The Astrophysical Journal, 679:31–51, 2008 May 20

PHOTOMETRIC REDSHIFTS AND SIGNAL-TO-NOISE RATIOS

V. E. Margoniner[1] and D. M. Wittman[1]
Received 2007 July 13; accepted 2007 December 3

ABSTRACT

We investigate the impact of photometric signal-to-noise ratio (S/N) on the precision of photometric redshifts in multiband imaging surveys using both simulations and real data. We simulate the optical four-band (*BVRz*) Deep Lens Survey (DLS) and use the publicly available Bayesian Photometric Redshift code BPZ. The simulations include a realistic range of magnitudes and colors and vary from infinite S/N to S/N = 5. The real data are from DLS photometry and two spectroscopic surveys and explore a range of S/Ns by adding noise to initially very high S/N photometry. Precision degrades steadily as S/N drops, both because of direct S/N effects and because lower S/N is linked to fainter galaxies with a weaker magnitude prior. If a simple S/N cut were used, S/N $\geq$ 17 in *R* (corresponding, in the DLS, to lower S/N in other bands) would be required to keep the scatter in $\Delta z \equiv (z_{\rm spec} - z_{\rm phot})/(1 + z_{\rm spec})$ to less than 0.1. However, cutting on ODDS (a measure of the peakiness of the probability density function provided by BPZ) greater than 0.4 provides roughly double the number of usable galaxies with the same $\sigma_{\Delta z}$. Ellipticals form the tightest $z_{\rm spec}$-$z_{\rm phot}$ relation, and cutting on elliptical type provides better precision than the ODDS > 0.9 cut, but this eliminates the vast majority of galaxies in a deep survey. In addition to being more efficient than a type cut, ODDS also has the advantages of working with all types of galaxies (although ellipticals are overrepresented) and of being a continuous parameter for which the severity of the cut can be adjusted as desired.

Subject headings: galaxies: distances and redshifts — galaxies: photometry — methods: data analysis — surveys

Online material: color figures

1. INTRODUCTION

Photometric redshifts (Connolly et al. 1995; Hogg et al. 1998; Benítez 2000) are of paramount importance for current and planned multiband imaging surveys. With photometric redshifts, surveys can inexpensively gather information about structure along the line of sight without resorting to expensive spectroscopic follow-up. Therefore, it is important to understand the systematic errors and limitations in this method. For example, Ma et al. (2006) and Huterer et al. (2006) have examined the required photometric redshift accuracy for surveys which plan to use weak lensing (cosmic shear) to constrain dark energy. For this application and for baryon acoustic oscillations (Zhan & Knox 2006), reducing photometric redshift errors is less important than knowing the error distribution accurately. Thus, careful attention must be paid to systematic differences between the photometric survey and the spectroscopic sample used to evaluate photometric redshift performance. For most surveys, photometric signal-to-noise ratio (S/N) is one of the systematic differences.

The most well-known test case for photometric redshifts is the blind test in the Hubble Deep Field North (HDFN) conducted by Hogg et al. (1998). The best methods then yielded $\sigma_{\Delta z} \sim 0.1$, where $\Delta z \equiv (z_{\rm spec} - z_{\rm phot})/(1 + z_{\rm spec})$, using *Hubble Space Telescope* (*HST*) photometry in the *UBVI* bands and ground *JHK* (Dickinson 1998). More recently, with improved photometry and spectral redshift classification, an accuracy of $\sigma\Delta z \sim 0.06$ is achieved over the redshift range 0–6 (Fernández-Soto et al. 1999, 2001; Benítez 2000). Ground-based surveys suffer from less precise photometry but usually do not have to deal with such a large redshift range. Ilbert et al. (2006) cite an accuracy of $\sigma_{\Delta z} = 0.029$ after clipping outliers with $\Delta z > 0.15$ (3.8% of the sample). Ilbert et al. (2006) also find a decrease in precision at fainter magnitudes but make no effort to separate the effects of S/N from the other effects operating on faint galaxies, such as a weaker magnitude prior and greater spectral energy distribution (SED) evolution. In this paper we examine the impact of these effects separately, focusing on photometric S/N. The quantitative results presented here are specific to the *BVRz* filter set used in the Deep Lens Survey (DLS; Wittman et al. 2002). More filters, covering a wider range in wavelengths, will do better (Abdalla et al. 2007). However, the trends with S/N are broadly applicable.

2. METHOD

We use the Bayesian Photometric Redshift code BPZ developed by Benítez (2000). We also tested the HyperZ code (Bolzonella et al. 2000) with additional priors roughly equivalent to the default BPZ priors and found similar performance. For clarity we present only the results from BPZ here. We did not test training-set methods, in which a spectroscopic and photometric training set is used to perform a fit or to train a neural network, for two reasons. First, training set methods are unlikely to be employed for surveys planning to push the photometric sample deeper than the spectroscopic sample. Second, the two methods seem to be roughly equivalent in performance on the data sets in which they have been compared (e.g., Hogg et al. 1998), so the trends presented here should be applicable to both methods.

We use the six SED templates from Benítez (2000): E, Sbc, Scd, Irr, SB3, and SB2, modified as described below. For the simulations, the same templates are used to simulate the photometry and to infer the photometric redshifts; there is no allowance for cosmic variance of the templates or "template noise." For the data, it is important that the templates reflect real SEDs. Therefore, we use the photometry of objects with spectroscopic redshifts to optimize the templates (Csabai et al. 2000; Benítez et al. 2004; Ilbert et al. 2006). Section 4.2 describes the procedure and shows the corrected templates. Clearly, even the optimized templates do not represent all types of SEDs in the universe. For both simulations and data, we start by demonstrating the performance with

[1] Department of Physics, University of California at Davis, 1 Shields Avenue, Davis, CA 95616.

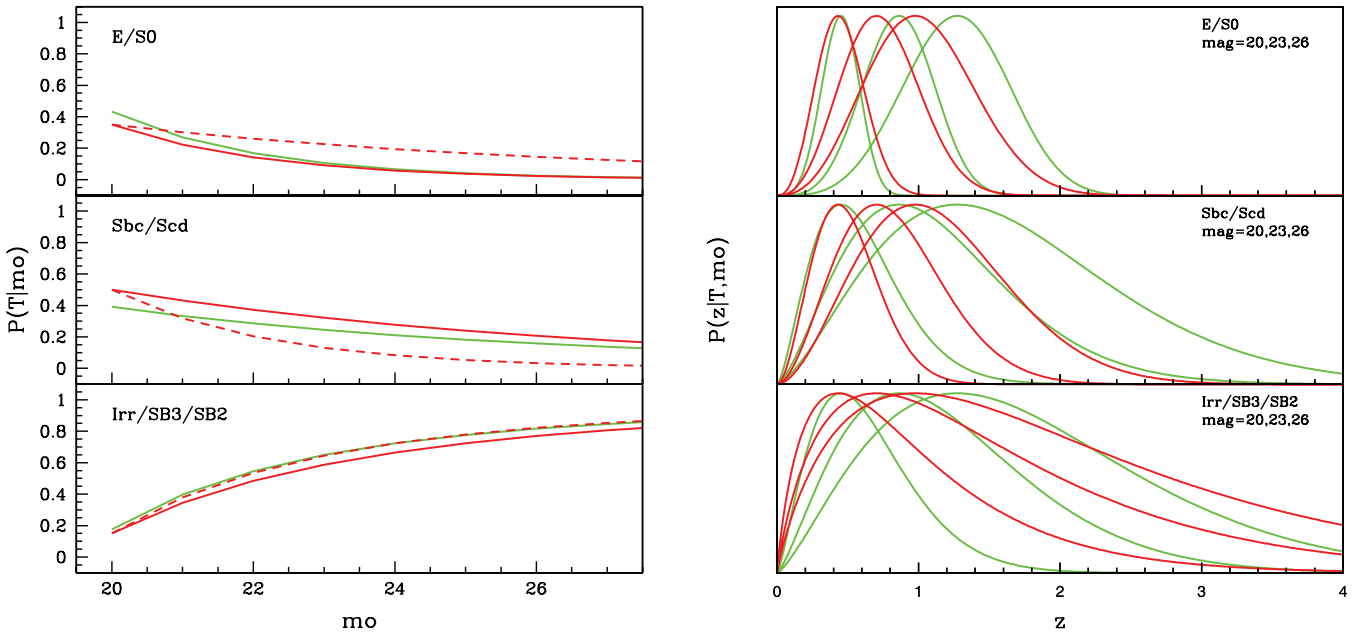

FIG. 1.—Priors used to populate the simulations. *Left*: $P(T|m_0)$ is the galaxy-type fraction as a function of magnitude. *Right*: $P(z|T, m_0)$ is the redshift distribution for galaxies of a given spectral type and magnitude for mag = 20, 23, and 26. Throughout this paper we use the priors indicated by the solid red lines (BPZ code). The dashed red lines represent the priors in BPZ (Benítez 2000), while the green lines indicate the priors derived by Ilbert et al. (2006).

as nearly perfect a data set as possible. After illustrating the best-case scenarios, we proceed to degrade the simulations and data to successively lower S/N, repeating the analysis for each step.

For each galaxy, we identify the peak of its redshift probability density function (PDF) as its *photometric redshift*, or $z_{\rm phot}$. This greatly simplifies the analysis and presentation of the results, at the cost of some precision. Specifically, "catastrophic outliers" will appear whose $z_{\rm phot}$ differs greatly from their true redshift. In many cases, this may be an artifact of not considering the full PDF, a point argued forcefully in the case of the HDF by Fernández-Soto et al. (2001, 2002). The full PDF may contain additional peaks or otherwise be broad enough to be consistent with the true redshift. In this paper we wish to focus on the trends with photometric S/N rather than the characterization of outliers. As seen in the tables and figures, the trends with S/N are not substantively changed if outliers are removed. Therefore, we judge this simplification to be acceptable. "Outlier" in this paper thus refers to the difference between $z_{\rm phot}$ and the true redshift without implying anything about the full PDF.

We do consider characteristics of the PDF when using BPZ's ODDS parameter. BPZ assumes a natural error (template noise) of $0.067(1+z)$ and defines ODDS as the fraction of the area enclosed by the PDF between $z_{\rm phot} \pm 0.067n(1+z_{\rm phot})$, where n is a user-settable parameter, which we set to 1. ODDS values close to unity indicate that most of the area under the redshift PDF is within $Z_B \pm 0.067(1+z)$. In this paper we present results both for the entirety of a given sample and after a cut of ODDS > 0.9, which eliminates many of the outliers. We also investigate the trade-off between the ODDS cut, the number of usable galaxies, and photometric redshift accuracy.

The error distributions are typically non-Gaussian, often highly so. The rms or standard deviation is extremely sensitive to even a few non-Gaussian events, so in the photometric redshift literature, results are usually quoted as an rms after excluding a certain (small) fraction of galaxies as catastrophic outliers. The fraction varies from paper to paper, making comparison difficult. The field of robust statistics suggests several less sensitive metrics of variation, such as the median or mean absolute deviation. However, outliers *should* be included in the performance analysis with some weight, because they will be included when using the entire photometric sample for science. We therefore clip conservatively, $|\Delta z| < 0.5$, to avoid overly optimistic results. This threshold is at least 5, and usually many more, times the clipped rms. We also present, in many cases, differential and cumulative distributions as well. To make the connection with forecasts for, say, weak-lensing tomography, we suggest these distributions be fit with double Gaussians. Gaussians are analytically tractable, and a double Gaussian can fit both the core and wings (but not truly catastrophic outliers).

3. SIMULATIONS

We simulate a mix of ellipticals, spirals, irregulars, and starburst galaxies (specifically, E, Sbc, Scd, Irr, SB3, and SB2 templates) following the priors for galaxy-type fraction as a function of magnitude, $P(T|m_0)$, and for the redshift distribution for galaxies of a given spectral type and magnitude, $P(z|T, m_0)$, that are used in BPZ. We found that in Table 1 of Benítez (2000), two numbers were inadvertently switched, but the numbers were correct in the publicly downloadable code. N. Benítez (2006, private communication) has confirmed that the table should read $k_t = 0.450$ for E/SO and $k_t = 0.147$ for Sbc/Scd. Figure 1 shows (*solid red lines*) the priors used in this paper (same as BPZ); also shown (*dashed red lines*) are the priors quoted in Benítez (2000) and (*green lines*) the Ilbert et al. (2006) priors.

In order to have a realistic galaxy luminosity function, N(mag), we start our simulations from R-band magnitudes of 87,260 objects detected in one of our $\sim 40' \times 40'$ DLS subfields (Wittman et al. 2002). The typical $BVRz$ magnitude distributions for the DLS are shown in Figure 2. We take this magnitude as the *true* (*≠observed*) R-band magnitude of a new object to be simulated. From the $P(T|m_0)$ prior we select a SED, and from $P(z|T, m_0)$ we choose a $z_{\rm input}$ redshift for the galaxy. The resulting "true" redshift distribution in the simulations is shown in Figure 3. This distribution has a larger tail to high redshift than usually found in the

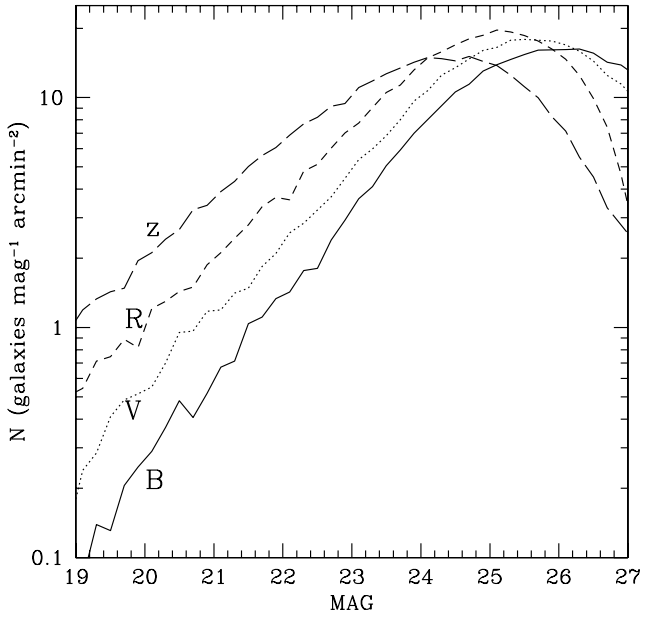

FIG. 2.—DLS $N(\mathrm{MAG_{iso}})$ for *BVRz*. [*See the electronic edition of the Journal for a color version of this figure.*]

literature (e.g., Le Fèvre et al. 2005) and can be approximately described as $z^2 \exp[-1(z/0.05)^{0.54}]$. Magnitudes (with or without noise) in any other photometric bands can then be computed. We use BPZ itself to compute synthetic colors, so there is no issue of minor differences in the *k*-corrections, priors, etc. We assume that there are only six SEDs of galaxies in the universe and make no attempt to introduce template noise in these simulations. We then perform three sets of simulations in the *BVRz* filter set of the DLS. In the first simulation (SIM1) we assume perfect, infinite S/N photometry. In the second set of simulations (SIM2) we successively degrade the S/N of the photometry but maintain the S/Ns of all galaxies in all four bands constant (same magnitude error for all galaxies in all four bands). In the third simulation (SIM3) we reproduce the S/N distribution and completeness of the DLS.

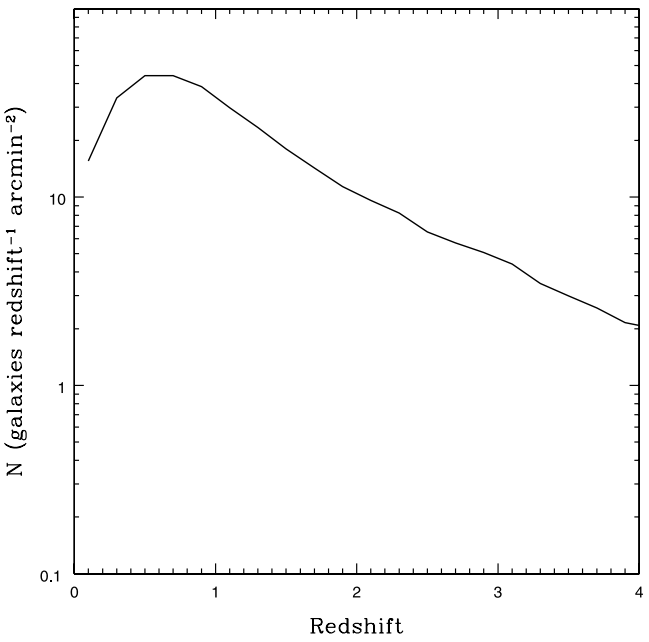

FIG. 3.—$N(z_{\mathrm{input}})$ for simulations.

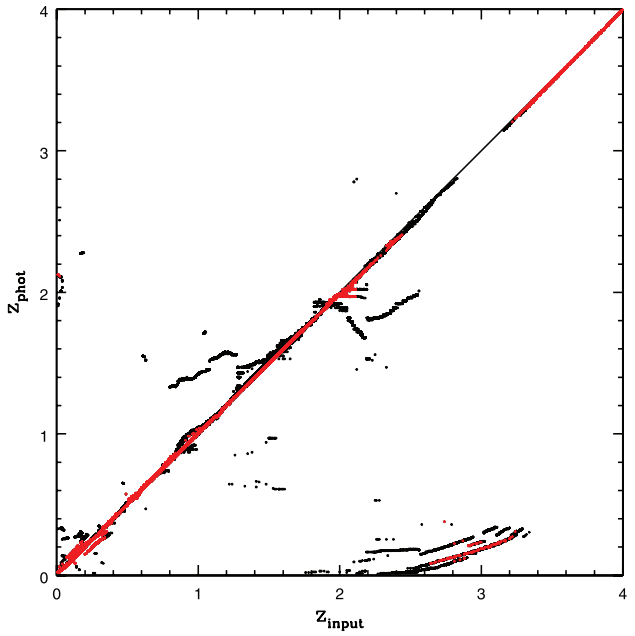

FIG. 4.—The z_{phot}-z_{spec} scatter plot for SIM1 (no photometry noise). Galaxies with ODDS > 0.9 are in red. See Table 1 for statistics.

3.1. *SIM1*

The first simulation (SIM1) has perfect photometry and represents the best possible case. The z_{phot}-z_{spec} scatter plot for this simple simulation is shown in Figure 4, and the distribution of $\Delta z \equiv (z_{\mathrm{spec}}\text{-}z_{\mathrm{phot}})/(1 + z_{\mathrm{spec}})$ is shown in Figure 5. Note that Figure 4 contains 87,260 objects, distributed in redshift according to Figure 3, and that the $z_{\mathrm{phot}} = z_{\mathrm{spec}}$ line is saturated with objects. It is clear from Figure 5 that the majority of objects have $|\Delta z| \sim 0.0$. Table 1 indicates (1) the S/N of the photometry (same in all bands), (2) the fraction of galaxies with $|\Delta z| < 0.5$, (3) the mean

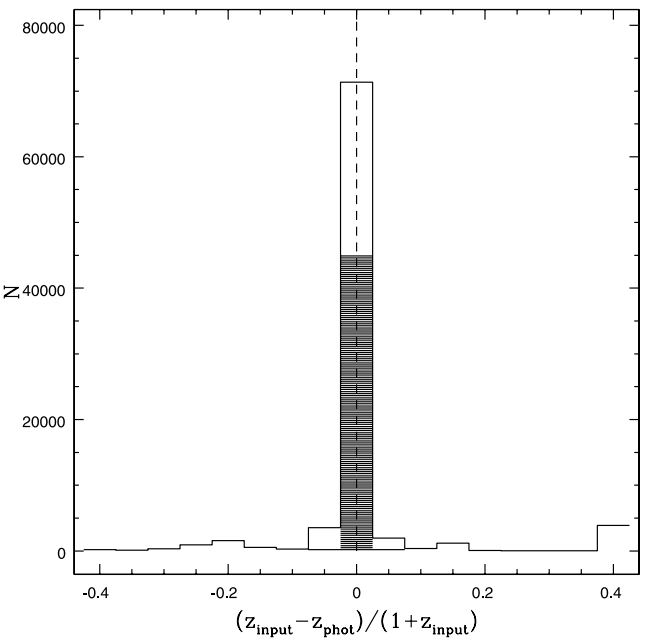

FIG. 5.—Histogram of Δz for the simulation in Fig. 4. The distribution of galaxies with ODDS > 0.9 is shaded. The outermost bins show the integrated counts of all objects with $|\Delta z| > 0.4$. [*See the electronic edition of the Journal for a color version of this figure.*]

TABLE 1
GALAXIES WITH FIXED S/N (SIM1 AND SIM2)

	$\lvert\Delta z\rvert \leq 0.5$			ODDS > 0.9			
S/N	$\lvert\Delta z\rvert \leq 0.5$/all (%)	$\bar{\Delta z}$	$\sigma(\Delta z)$	ODDS > 0.9/all (%)	$(\lvert\Delta z\rvert \leq 0.5, \text{ODDS} > 0.9)$/ODDS > 0.9 (%)	$\bar{\Delta z}$	$\sigma(\Delta z)$
Inf (SIM1).....	95.3	−0.008	0.051	53.4	99.7	−0.000	0.006
250 (SIM2).....	96.0	−0.005	0.042	64.2	99.8	−0.000	0.009
100 (SIM2).....	95.5	−0.007	0.049	60.7	99.7	−0.000	0.012
60 (SIM2).....	94.7	−0.010	0.062	54.4	99.7	−0.001	0.015
30 (SIM2).....	92.5	−0.014	0.085	40.6	99.9	−0.002	0.020
10 (SIM2).....	87.6	−0.007	0.121	6.4	100.0	−0.001	0.023
5 (SIM2).....	84.7	0.019	0.151	1.2	99.9	0.001	0.012

Δz for galaxies with $|\Delta z| < 0.5$, (4) the rms in Δz for galaxies with $|\Delta z| < 0.5$, (5) the fraction of objects with ODDS > 0.9, (6) the fraction of objects with ODDS > 0.9 and $|\Delta z| < 0.5$, (7) the mean Δz, and (8) the rms in Δz for these galaxies.

There are still catastrophic outliers, despite this being the best possible case in terms of noise and perfectly known templates and priors. This is due to color-space degeneracies and the representation of each galaxy as a single point rather than using the full PDF, as discussed in § 2. Even though the magnitude prior is perfectly known here, it is not strong enough to "break" the color-space degeneracy in the sense of choosing the correct peak for each galaxy. Rather, it (in principle) modifies the PDFs so that they are not inconsistent with the redshifts. As our purpose is only to establish SIM1 as a baseline for investigating the impact of

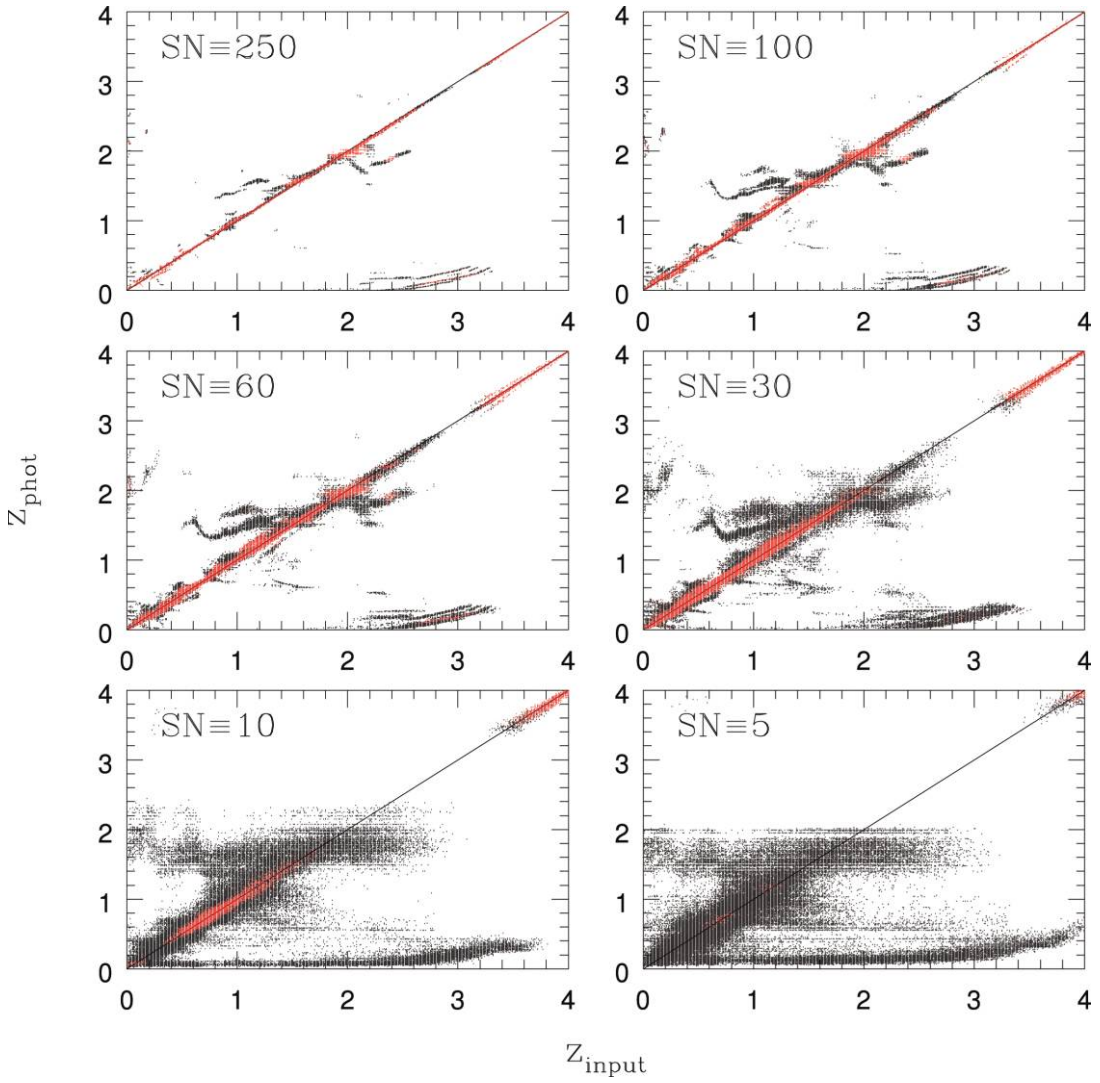

FIG. 6.—The $z_{\rm phot}$-$z_{\rm spec}$ scatter plot for simulations with realistic magnitude and redshift distributions but uniform and progressively greater photometry noise (SIM2). In each panel all galaxies have the same S/N in *BVRz*. *Left to right, top to bottom*: S/N = 250, 100, 60, 30, 10, and 5. See Table 1 for statistics.

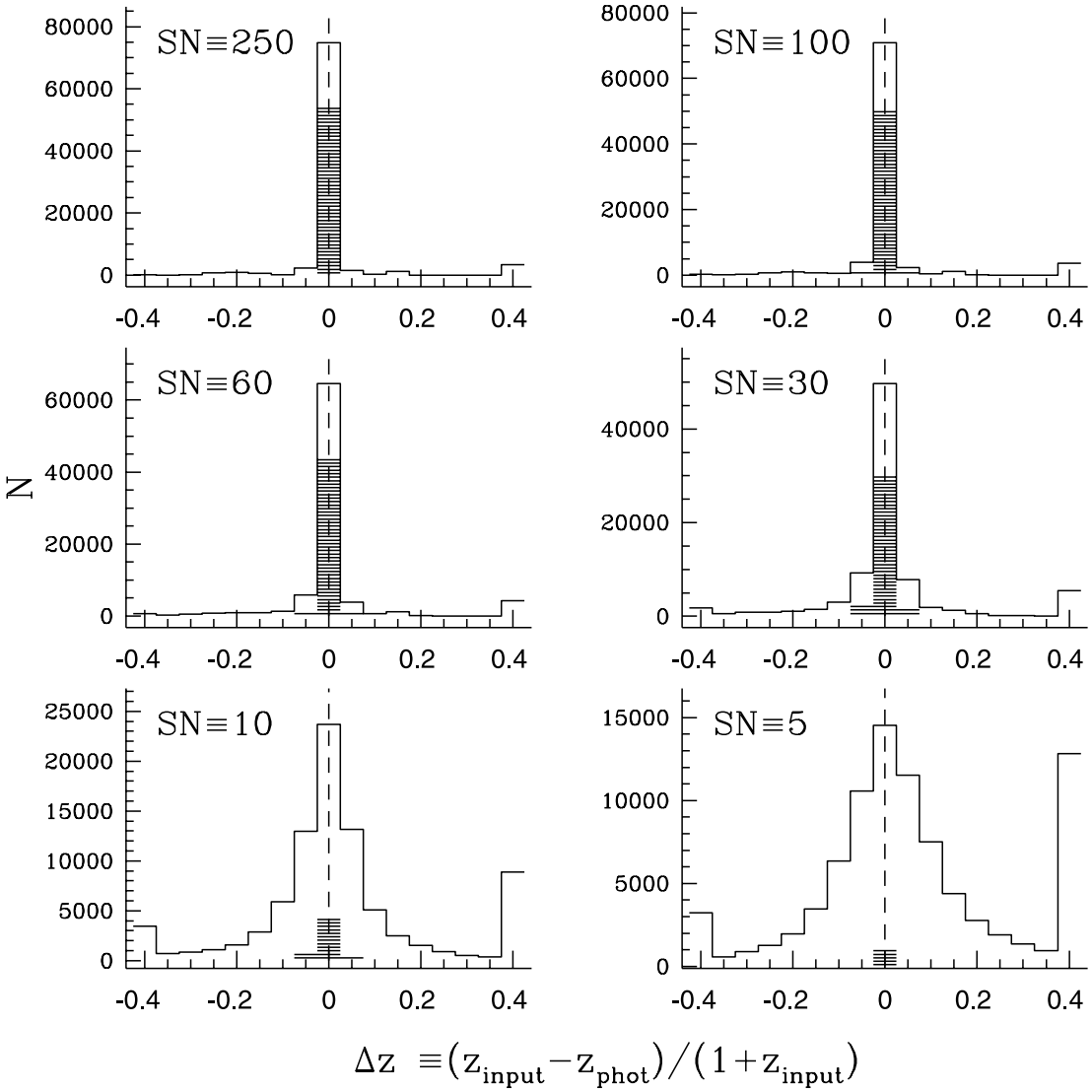

FIG. 7.—Histogram of Δz for objects shown in Fig. 6 (SIM2). The distribution of galaxies with ODDS > 0.9 is shaded. [*See the electronic edition of the Journal for a color version of this figure.*]

photometric S/N, we do not pursue PDF analysis further in this paper.

3.2. *SIM2*

In the second set of simulations (SIM2) we degrade the initially perfect photometry in SIM1 successively to an S/N of 250 ($R \sim 20.5$ mag in the DLS and the magnitude limit of the spectroscopic sample presented in § 4), 100, 60, 30, 10, and 5 and repeat the analysis at each step. In these unrealistic simulations all galaxies have the same photometric S/N in all bands. The scatter plots are shown in Figure 6, and Δz distributions are shown in Figure 7. We also present the cumulative fraction of objects with Δz smaller than a given value, as a function of Δz (Fig. 8). This plot has several advantages. First, multiple simulations can be overplotted without obscuration. Second, the asymmetry in the distribution of Δz is easily read off by looking at the fraction with $\Delta z < 0$ (*dashed vertical line*). Third, the fraction of outliers can also be directly read off the plot at any Δz. The left panel of Figure 8 shows the cumulative fraction for all objects, while the right panel shows ODDS > 0.9 galaxies. The number of galaxies in the right panel is smaller than the number in the left panel (see Table 1), but the accuracy of photo-zs is clearly better.

Because all realizations of SIM2 have the redshift distribution shown in Figure 3, even if all galaxies have colors measured at very high S/N, some objects will have degenerate colors and the sample will contain some fraction of catastrophic outliers. Spectroscopic samples typically have a much lower mean redshift than these simulations, so catastrophic outliers are likely to be underrepresented in direct $z_{\rm phot}$-$z_{\rm spec}$ comparisons, if the full photometric sample is very deep.

Table 1 presents the statistics for the SIM2 objects shown in Figures 6–8. Clearly, the precision of photometric redshifts is a strong function of photometric S/N. The BPZ ODDS parameter is very effective at removing outliers, and almost 100% of the objects with ODDS > 0.9 have $|\Delta z| < 0.1$ regardless of S/N (Fig. 8, *right*). However, the fraction of objects with ODDS > 0.9 decreases dramatically with decreasing S/N.

Performance is, counterintuitively, slightly worse for the infinite-S/N galaxies in SIM1 than for the high-S/N galaxies in SIM2. This is because PDFs are unrealistically narrow when there is no noise in color space, and priors only serve to smear them out. (The SIM1 PDFs were not true delta functions because BPZ was run with a small nominal error to avoid division by zero.) This introduces ambiguities where there were near-degeneracies in color space.

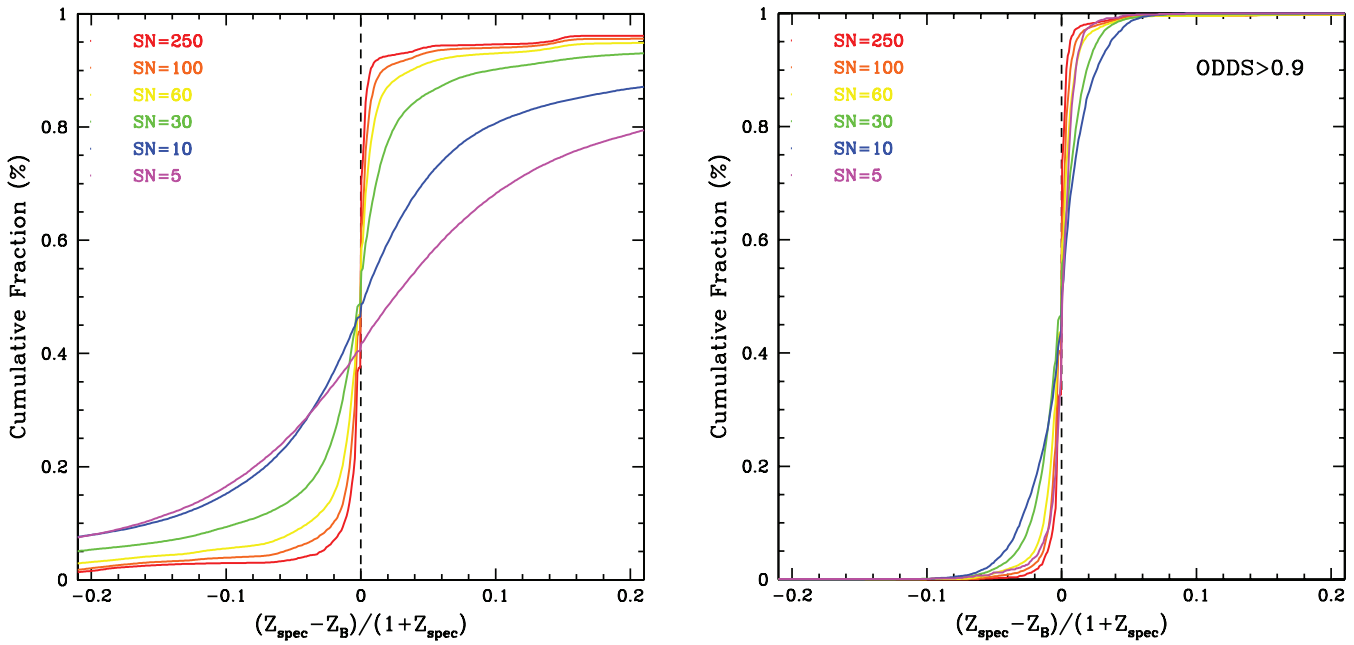

FIG. 8.—Cumulative fraction of objects with Δz smaller than a given value. The red line indicates the simulation in which all galaxies have been set to have S/N = 250 in all *BVRz*, orange indicates a simulation with S/N = 100, and so on. The left panel shows all the galaxies, and the right panel shows galaxies with ODDS > 0.9. Note that only 6.4% and 1.2%, respectively, of objects with S/N = 10 and 5 have ODDS > 0.9.

In the presence of photometry noise, the near-degeneracies already present ambiguity, and the priors help choose the correct option. We further illustrate this effect by looking at the results using a flat prior (dubbed the "ML" case, for "maximum likelihood"). Because of this effect, ML slightly outperforms the prior in SIM1 ($\sigma_{\Delta z} = 0.170$ vs. 0.190 with no outlier clipping). At S/N = 250 in SIM2, the situation has reversed: with no clipping, $\sigma_{\Delta z} = 0.190$ for ML versus 0.156 for the prior. This is mostly due to catastrophic outliers eliminated by the prior, because with clipping at $|\Delta z| < 0.5$, ML is not worse than the prior. In real life, one cannot clip in this way, so we conclude that the prior is helpful even at this high S/N. As the S/N in SIM2 drops, the ML performance steadily drops relative to that of the prior. At S/N = 10, the unclipped $\sigma_{\Delta z}$ for ML is more than double that for the prior, 0.659 versus 0.292, and even the clipped value is significantly worse, 0.143 versus 0.121 (the uncertainties in all these values are ±0.001 or better).

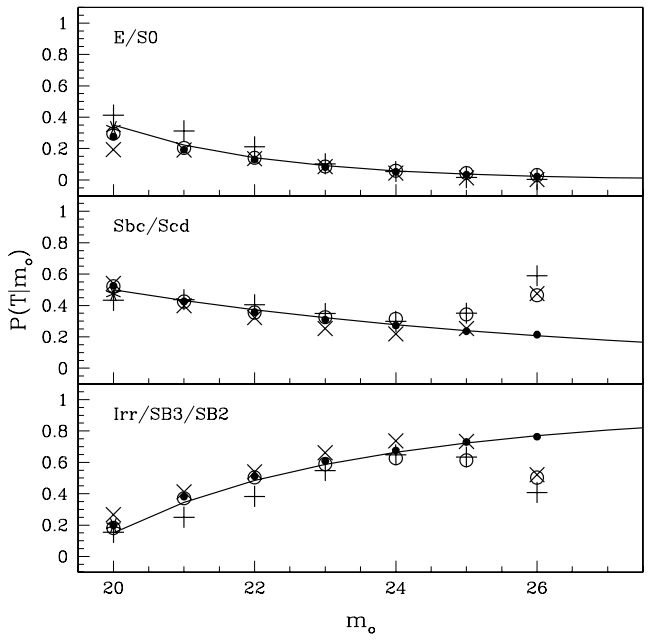

FIG. 9.—Galaxy-type fraction as a function of magnitude, $P(T|m_0)$. The curves indicate the BPZ priors used in the simulations of § 3 and shown in Fig. 1. The plus signs and crosses indicate the fraction of galaxies classified by BPZ as E, Sbc/Scd, or Im/SB3/SB2 in two DLS fields of $40' \times 40'$ each. The plus-sign field, with a higher fraction of ellipticals, contains the galaxy cluster Abell 781, while the crosses represent a more typical "blank" field. The simulation input distribution is indicated by filled circles, which by definition agree with the curves, while the open circles indicate the BPZ type classification of these objects. The asterisks represent the SHeLS spectroscopic sample. [*See the electronic edition of the Journal for a color version of this figure.*]

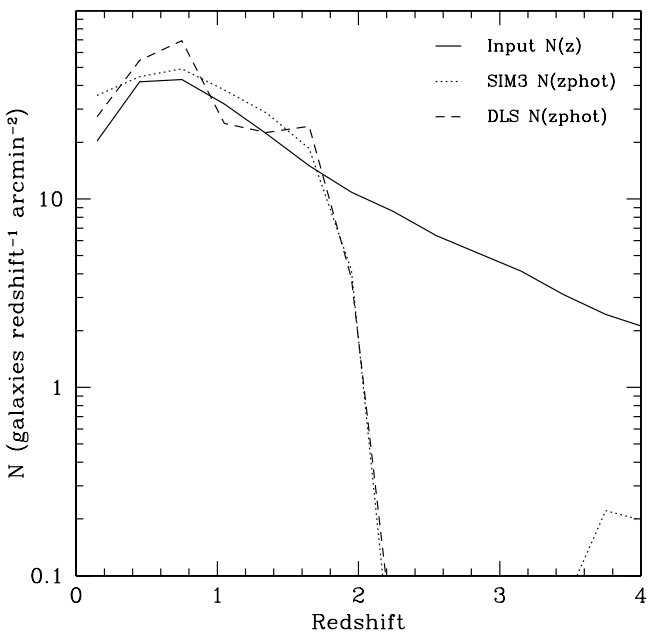

FIG. 10.—Photometric redshift distributions for the DLS and SIM3. The input $N(z)$ for the simulations is shown for comparison. [*See the electronic edition of the Journal for a color version of this figure.*]

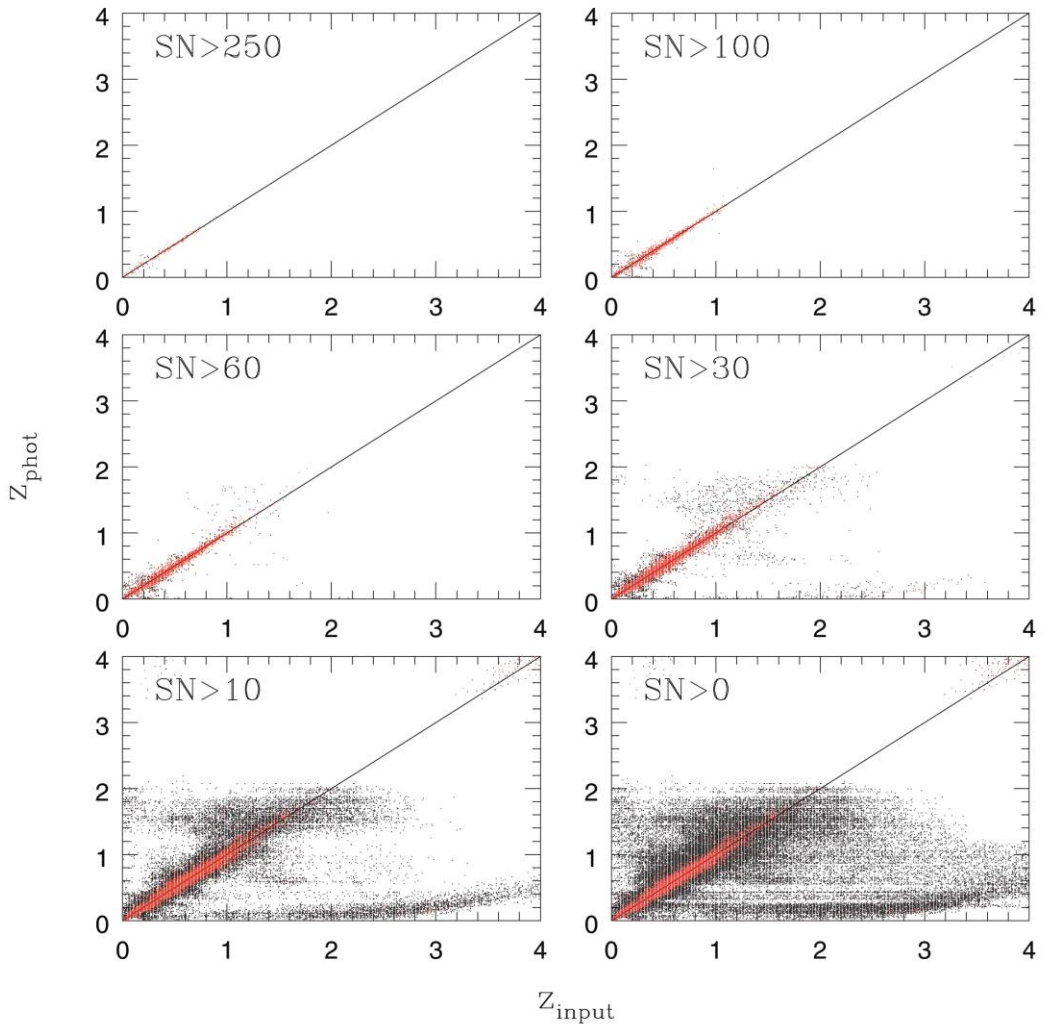

FIG. 11.—The $z_{\rm phot}$-$z_{\rm spec}$ scatter plot for simulations with realistic magnitude, redshift, and S/N distributions (SIM3). The top left panel shows all the galaxies with very high S/N (small number density and mostly at low redshift), while the bottom right panel includes all galaxies in the simulation. *Left to right, top to bottom*: galaxies with S/N $>$ 250, >100, >60, >30, >10, and all.

3.3. *SIM3*

The third simulation has the same S/N distribution and completeness as the DLS data. Again, the priors used assure that the galaxy-type mixture and redshift distribution should be close to the real universe. The idea is to measure how well we can recover true $z_{\rm input}$ redshifts for a realistic photometric data set. This simulation is still optimistic because no template noise is added; we derive colors from the same six templates used in the determination of photometric redshifts. The effect of template noise is presented in the real data analysis in § 4.

As a sanity check we compare the *BVz* magnitude distributions of our SIM3 simulation with the observed N(mag) and find good agreement. The R-magnitude distribution is by definition the same within the added photometric noise. We also compare the distribution of BPZ galaxy types in DLS fields with the one derived from the SIM3 simulation and find very good agreement. Figure 9 shows the galaxy-type fraction as a function of magnitude for two $40' \times 40'$ DLS fields. The field with the higher fraction of ellipticals contains the richness class 2 galaxy cluster Abell 781 at $z = 0.298$ (*plus signs*), and the other is a more typical "blank" field (*crosses*). The simulation input distribution is indicated by filled circles, which by definition agree with the red line, and the output BPZ types are indicated by open circles. SIM3 and the data show the same magnitude dependence.

A third sanity check is a comparison between the redshift distribution derived in SIM3 and $N(z)$ for the entire DLS. Figure 10 shows both distributions and also the input redshift distribution used in the simulations (same as Fig. 3). The agreement is pretty good. The mean density of galaxies with photometric redshifts of any quality is 47 arcmin^{-2}, and 11% of those objects have ODDS > 0.9.

The photometric redshift performance on SIM3 is shown in Figures 11–13, just as in Figures 6–8 for SIM2. The summary statistics for SIM3 are presented in Table 2. As in SIM2, the precision of the photometric redshifts is a strong function of S/N, and ODDS does a good job of cleaning up, at the cost of losing many low-S/N galaxies.

There are two notable differences with SIM2. First, in SIM3 there is a realistically strong correlation between high S/N and bright magnitudes. A bright magnitude implies a strong prior (most bright galaxies are at low redshift), whereas a faint galaxy

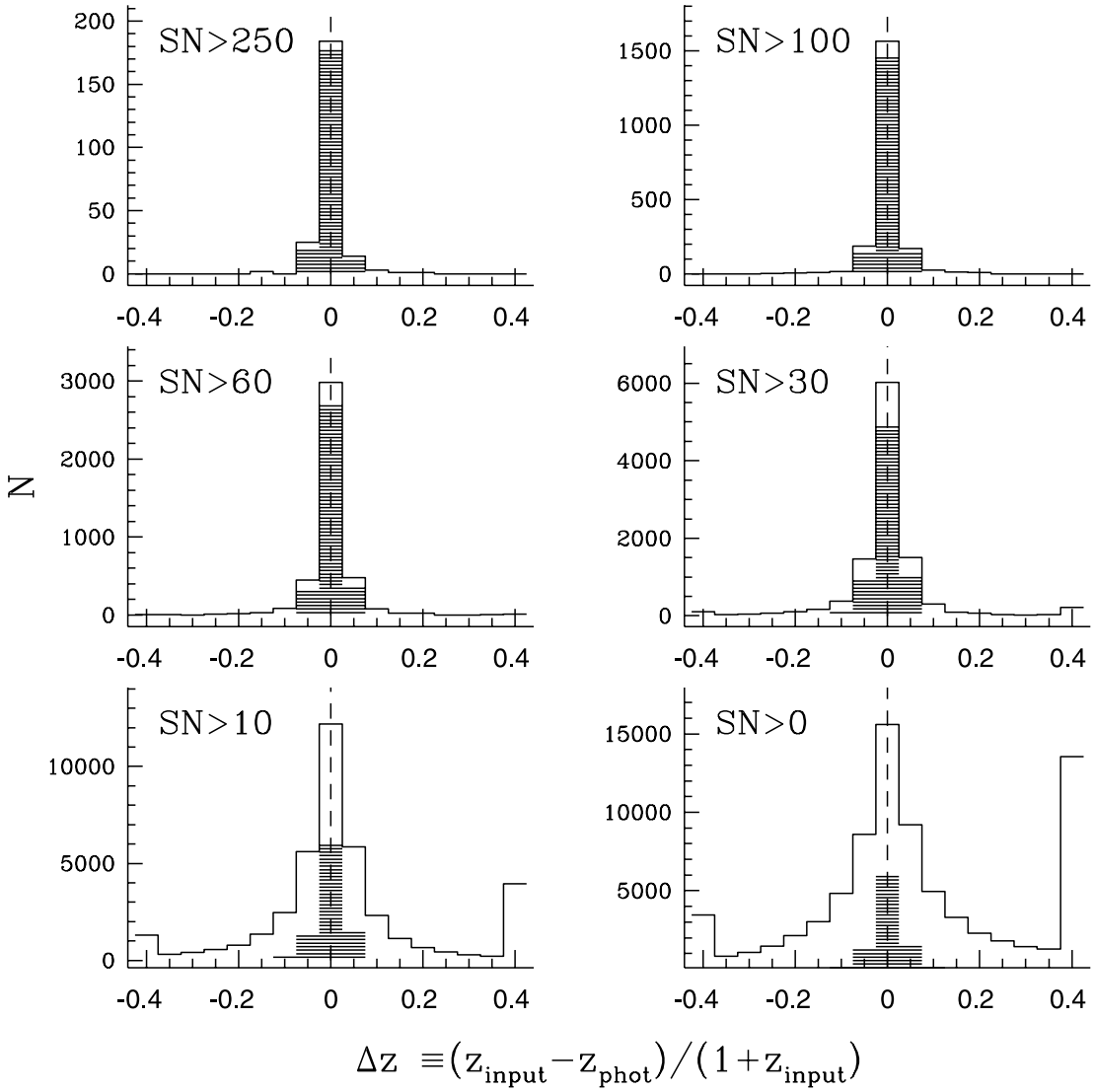

FIG. 12.—Histogram of Δz for objects shown in Fig. 11. The distribution of galaxies with ODDS > 0.9 is shaded. [*See the electronic edition of the Journal for a color version of this figure.*]

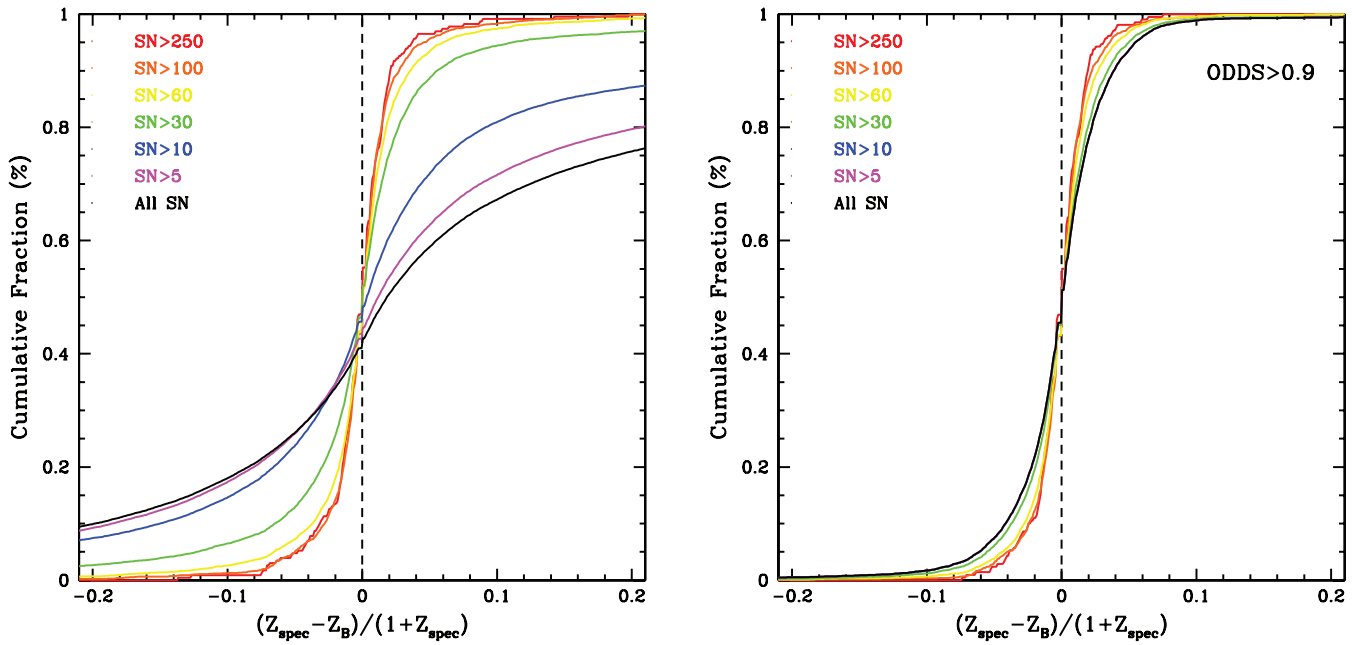

FIG. 13.—Cumulative fraction of objects with Δz smaller than a given value for simulations with realistic magnitude and redshift distributions (SIM3). The red line shows the cumulative fraction for all objects with S/N(R) > 250, orange shows the fraction for all objects with S/N(R) > 100 [including those with S/N(R) > 250], and so on. The left panel shows all the galaxies, and the right panel shows galaxies with ODDS > 0.9.

TABLE 2
GALAXIES IN DLS-LIKE SIMULATIONS (SIM3)

	$\lvert\Delta z\rvert \leq 0.5$			ODDS > 0.9			
S/N (R)	$\lvert\Delta z\rvert \leq 0.5$/all (%)	$\bar{\Delta} z$	$\sigma(\Delta z)$	ODDS > 0.9/all (%)	($\lvert\Delta z\rvert \leq 0.5$, ODDS > 0.9)/ODDS > 0.9 (%)	$\bar{\Delta} z$	$\sigma(\Delta z)$
>250	100.0	−0.001	0.031	90.9	100.0	−0.001	0.021
>100	100.0	0.001	0.037	89.4	100.0	0.000	0.026
>60	99.7	−0.000	0.050	82.6	100.0	0.000	0.030
>30	97.8	−0.004	0.076	67.0	99.6	−0.001	0.036
>10	89.3	−0.004	0.125	23.5	99.3	−0.001	0.040
>5	85.0	0.008	0.154	14.1	99.3	−0.001	0.040
All	83.6	0.018	0.170	11.9	99.3	−0.001	0.040

has a weak prior (it could be at any redshift). The high-S/N galaxies in SIM2 were (artificially) at all magnitudes and therefore had generally looser priors. Therefore, the highest S/N galaxies in SIM3 do better than those in SIM2. We can see the effect of the tight priors directly by comparing the S/N = 250 line of Table 1 ($\sigma_{\Delta z} = 0.042$ after clipping 4% which had $|\Delta z| > 0.5$) with that of Table 2 ($\sigma_{\Delta z} = 0.031$ with no need to clip any outliers). This difference vanishes when low-S/N galaxies from SIM3 are included.

In fact, the S/N = 5 galaxies in SIM2 outperform the S/N > 5 galaxies in SIM3, despite the latter cut being only a lower bound. This is due to the second salient difference between SIM2 and SIM3: a given S/N in SIM2 describes *each* galaxy in *each* band. In SIM3 the S/N varies with filter in a realistic way, and the cut applies to the *R* band. Most galaxies will have lower S/N in other bands. For S/N = 30 in *R*, the median S/N in *B*, *V*, and *z* over the whole sample is 10, 18, and 10, respectively.

What S/N is required for good photometric redshift performance? First, consider performance without any ODDS cut. At each step in Table 2 from S/N > 100 to >10, there is a 30%–50% increase in $\sigma_{\Delta z}$, so there is no natural break point. The $\sigma_{\Delta z}$ appears to stop this dramatic growth when stepping down from S/N > 10 to >5, but this is likely an artifact of clipping at $|\Delta z| > 0.5$, which is roughly 3 times the clipped rms at that point. Even at S/N > 10, $\sigma_{\Delta z}$ may be artificially low due to clipping, as more than 10% of galaxies were clipped. Most survey users would find the precision offered by the S/N > 30 cut acceptable but that of the S/N > 10 cut unacceptable. If we set $\sigma_{\Delta z} = 0.1$ as the limit of acceptability, we find that an S/N cut at 17 is required.

Now consider using the ODDS cut at 0.9. The value of $\sigma_{\Delta z}$ is always 0.04 or less, regardless of S/N. We suspect that for a given $\sigma_{\Delta z}$, the ODDS cut will provide more galaxies than the S/N cut, because ODDS responds to the properties of the color space, as well as to S/N. For example, high-precision S/N is not required if the galaxy is in a distinctive region of color space. In addition, ODDS can take proper account of different S/N in different bands, which a simple S/N cut in *R* does not. We investigate this by finding the ODDS cut which yields the same $\sigma_{\Delta z}$ as the S/N > 30 cut (0.076). We find that ODDS > 0.57 is required, which yields 30% of all detected galaxies versus the 13% yielded by the S/N cut.

We repeat this procedure for $\sigma_{\Delta z} = 0.1$. The required ODDS cut is >0.40, yielding 45% of all detected galaxies, while the required S/N cut at 17 yields only 26% of detected galaxies.

These fractions can all be read off Figure 14, which summarizes the results from SIM3. The left panels in Figure 14 show (1) the

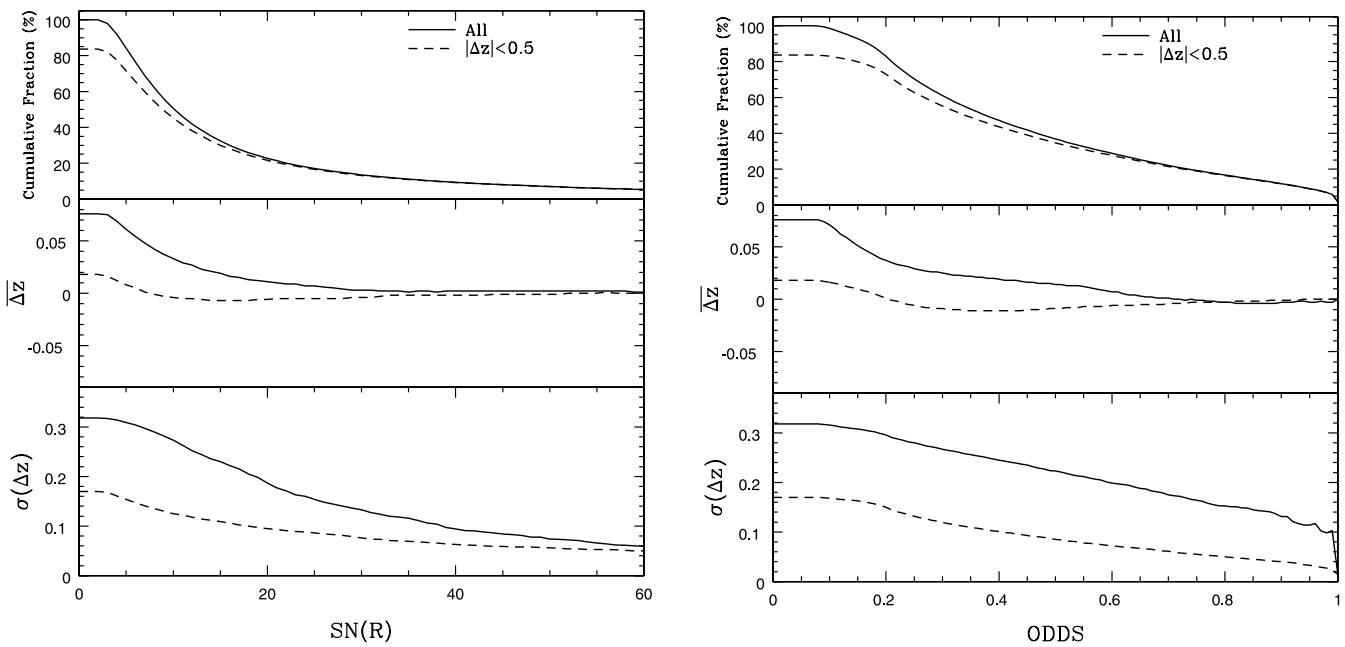

FIG. 14.—*Left*: Cumulative fraction of objects with S/N greater than a given value, mean Δz, and $\sigma_{\Delta z}$. *Right*: Same as left, but for objects with ODDS greater than a given value. The solid lines indicate all objects, and the dashed lines show $|\Delta z| < 0.5$.

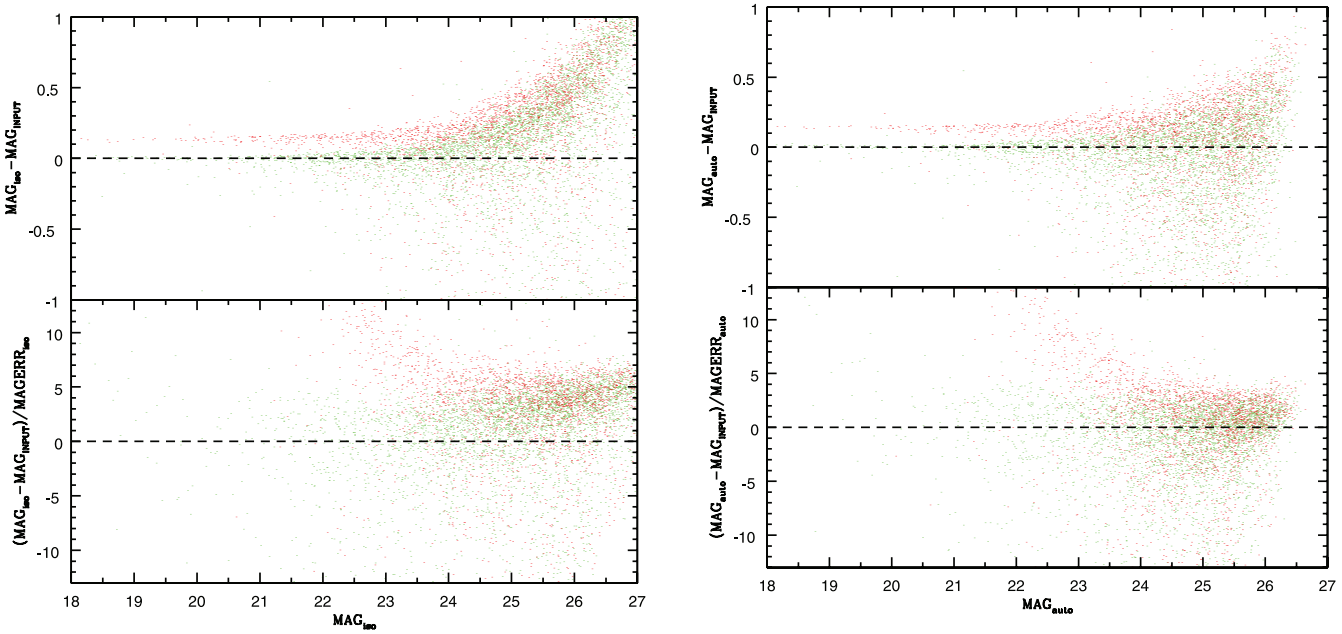

FIG. 15.—Magnitude errors of synthetic de Vaucouleurs (*red symbols*) and exponential disk (*green symbols*) galaxies added to DLS *R*-band images. The left panels show the SExtractor MAG_{iso} results, while the right panels indicate comparisons to SExtractor MAG_{auto}. *Top panels*: $\Delta MAG \equiv MAG - MAG_{input}$ vs. MAG. *Bottom panels*: $\Delta MAG/MAGerr$ vs. MAG.

cumulative fraction of objects with S/N greater than a given value, (2) the mean Δz, and (3) the value of $\sigma_{\Delta z}$ for these objects. The right panels are the same but for a cut in ODDS.

In short, we recommend an ODDS cut. We recognize that an ODDS cut is not easy to incorporate into survey forecasts of the number of usable galaxies. Detailed simulations for a given filter set and depth as a function of wavelength must be performed. However, we hope that the above numbers can serve as a rough guide for translation between photometric redshift precision, S/N threshold, and number of usable galaxies.

4. DATA

We take photometric data from the DLS *BVRz* full-depth images in fields with spectroscopic redshifts from the Smithsonian Hectospec Lensing Survey (SHeLS; Geller et al. 2005) and the Caltech Faint Galaxy Redshift Survey (CFGRS; Cohen et al. 1999). Here, by definition, template noise is present. In §§ 4.3 and 4.4 we present the spectroscopic data and the photometric redshift accuracy for these two samples, but before that we present our methodology for color measurement (§ 4.1) and template optimization (§ 4.2).

4.1. *Measuring Colors*

We performed simulations to determine the best photometry method in the face of different point-spread function (PSF) sizes in the different filters. We added galaxies with de Vaucouleurs (elliptical) and exponential disk (spirals) light profiles to the DLS *BVRz* data using standard IRAF `artdata` routines, ran SExtractor (Bertin & Arnouts 1996), and measured colors with the various types of magnitudes offered by SExtractor, in both single- and dual-image modes. In dual-image mode detection was in *R* because it is deeper than the other filters; very few galaxies not detected in *R* are detectable in any other filter. None of the magnitude types built into SExtractor are specifically designed to yield good colors in the presence of filter-dependent PSF size, but MAG_{auto} and MAG_{iso} in dual-image mode yielded acceptable results, in the sense that the color systematics are smaller than the photometric zero-point errors and therefore do not limit the photometry accuracy.

Figure 15 shows the magnitude errors for galaxies added to the *R* images. The *B*, *V*, and *z* results are qualitatively the same, but because of differences in S/N and PSF there is a shift in the magnitude axis and slightly different scatter. The left panels show the results using MAG_{iso}, and the right panels show MAG_{auto}. The top panels show the difference between measured MAG and MAG_{input}. De Vaucoleurs galaxies are measured to be $\sim$0.15 mag fainter than their true magnitudes by both MAG_{iso} and MAG_{auto}. The bottom panels show the distribution of $(MAG - MAG_{input})/MAGerr$ as a function of magnitude. MAG_{auto} gives better results for magnitudes, but we show below that MAG_{iso} gives better results for colors, which are the crucial quantities for photometric redshifts. The same pattern was noted by Benítez et al. (2004) for a very different data set.

Figures 16 and 17 show the distribution of color errors. Again, MAG_{iso} is on the left and MAG_{auto} on the right. The systematic magnitude errors tend to cancel when considering colors, and MAG_{iso} is now slightly better. It is important to note that the errors in magnitude are not driven by faint galaxies, and that in fact the discrepancies between real and estimated colors are significantly worse for bright objects.

In summary, MAG_{iso} gives slightly more precise colors at a given magnitude. This translates to more galaxies being detected above a given S/N threshold, providing another benefit. However, for either MAG_{auto} or MAG_{iso}, the error estimates provided by SExtractor are optimistic, especially at the bright end where the random error from sky noise is relatively unimportant. The solid lines in Figure 18 show the cumulative fraction of objects as a function of magnitude and color error, normalized by the nominal error from SExtractor. Much less than 68% (95%) of the galaxies have actual errors within the nominal 1 (2) σ magnitude error. Actual color errors are closer to nominal, but still optimistic. (A caveat is that, unlike most real galaxies, the simulated galaxies had zero color.) From this analysis we determine an ad hoc correction

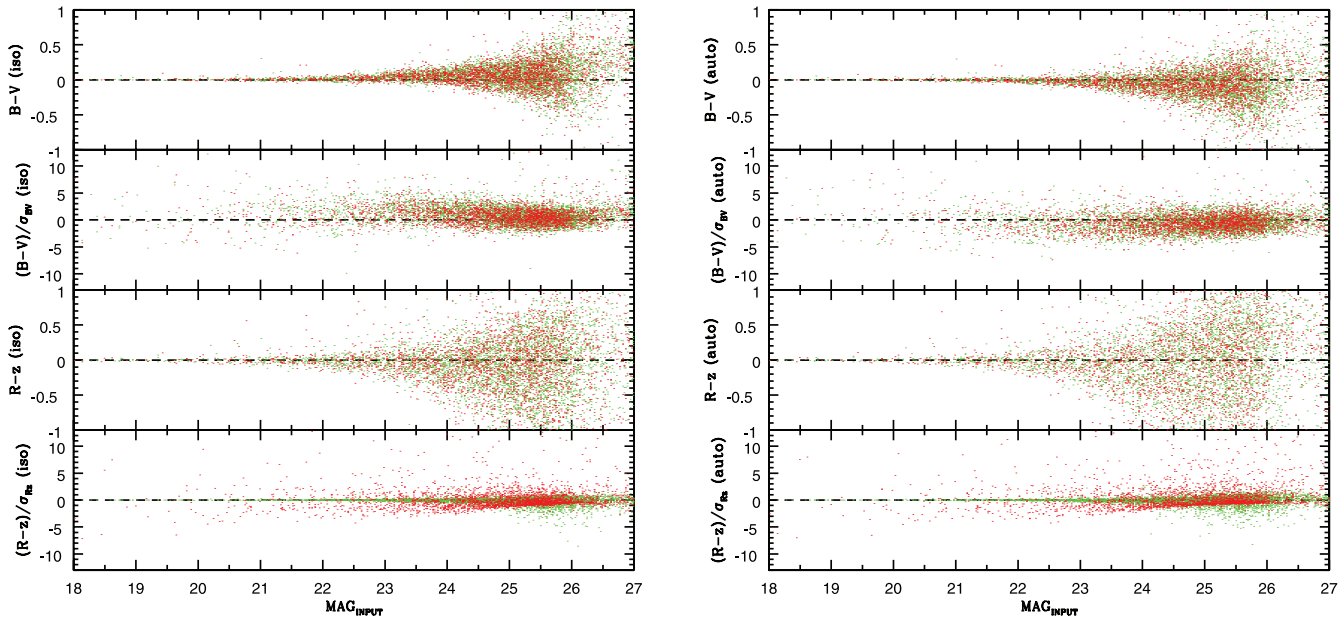

FIG. 16.—Color errors of synthetic de Vaucouleurs (*red symbols*) and exponential disk (*green symbols*) galaxies added to DLS images. The left panels show the SExtractor MAG_{iso} results, while the right panels indicate comparisons to SExtractor MAG_{auto}. The four panels in each column show different color combinations.

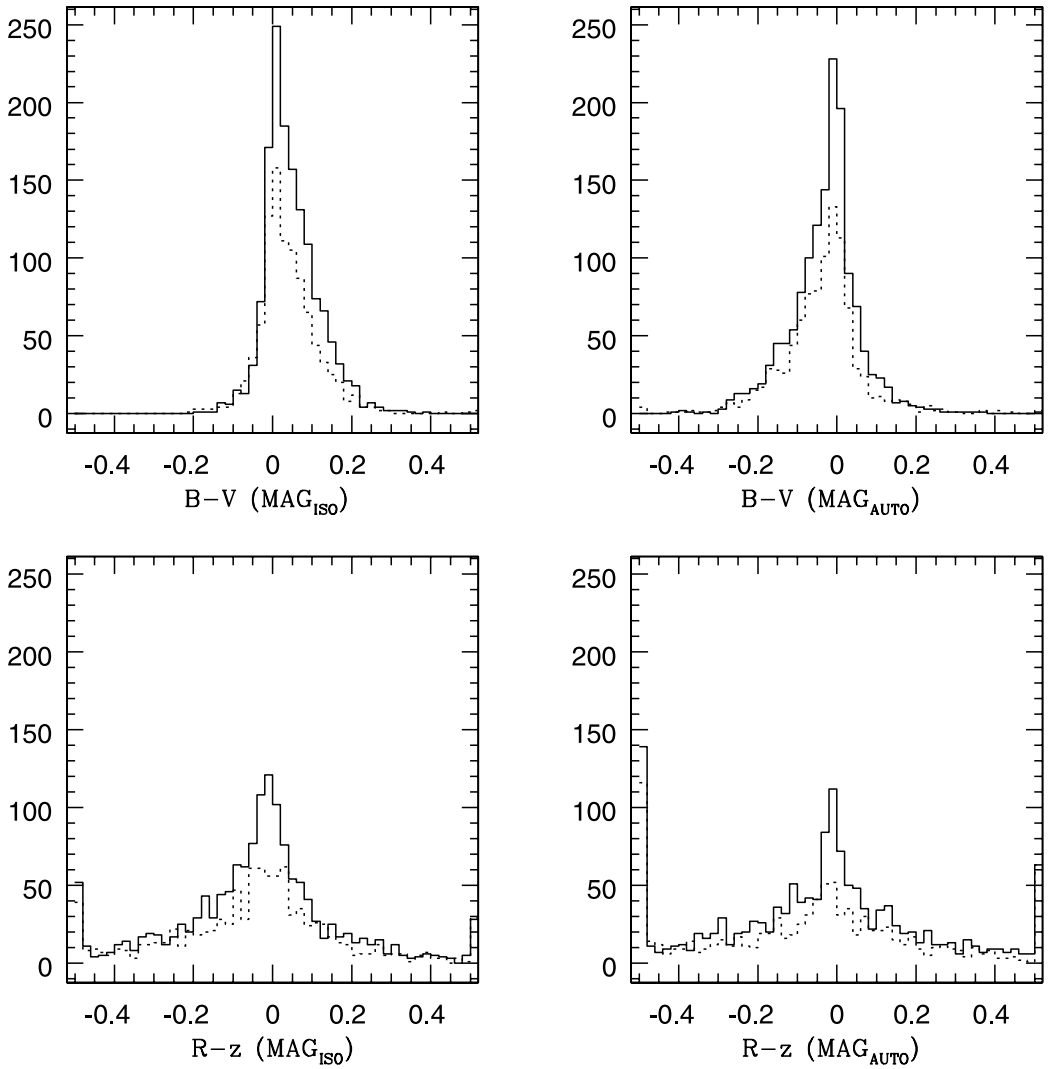

FIG. 17.—Distribution of colors derived from MAG_{iso} and MAG_{auto} for zero-color synthetic de Vaucouleurs (*dotted lines*) and exponential disk (*solid lines*) galaxies added to DLS data. Here we show galaxies brighter than 24.5 mag, which corresponds to S/N $\sim$ 10 in *BVR* but goes down to S/N $\sim$ 3 in *z*. The edge bins indicate the number of objects out of the limits of the plot. [*See the electronic edition of the Journal for a color version of this figure.*]

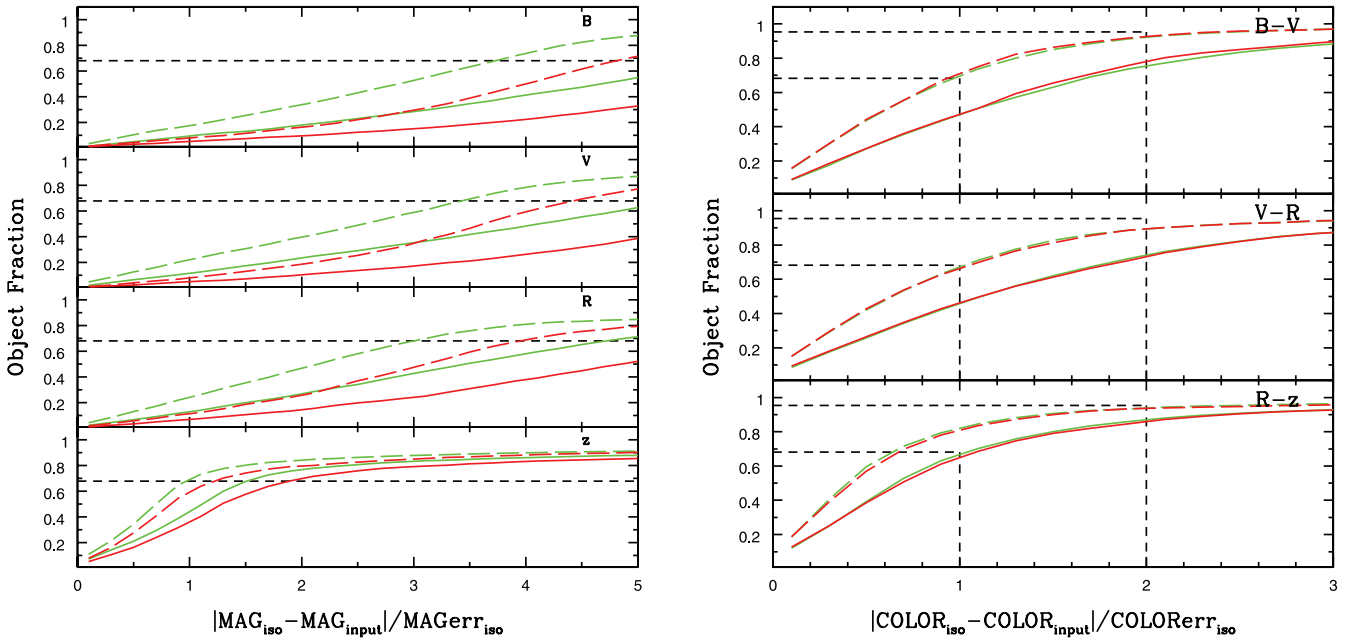

FIG. 18.—Cumulative fraction of objects as function of $\Delta_{MAG}/MAGerr_{iso}$ and $\Delta_{COLOR}/COLORerr_{iso}$. The red lines represent galaxies with a de Vaucouleurs light profile, and the green lines represent galaxies with an exponential disk. The dashed lines indicate the cumulative fraction after an ad hoc increase in the measured magnitude errors. The augmented errors guarantee that ~68% (95%) of the galaxies have colors within 1 (2) σ. A much larger increase would be needed in order to have ~68% (95%) of galaxies with measured magnitudes within 1 (2) σ.

to the magnitude errors estimated by SExtractor: we first multiply $MAGerr_{iso}$ by 1.5 and then add in quadrature an error of 0.02 mag. The dashed lines in both panels of Figure 18 show the results of this correction. This single correction puts the 68th and 95th percentiles of all the color distributions in the correct place, with the exception of the 68th percentile of $R - z$ color. This adjustment to the magnitude errors should, in principle, depend on galaxy color, but we found that variations about this correction made little difference in the results.

We performed all the real data photometric redshift tests in this paper with both MAG_{iso} and MAG_{auto}. The differences in the results were very minor, except that more galaxies were detected at a given S/N with MAG_{iso}, and about 20% more survived the ODDS cut with MAG_{iso}. We therefore adopt MAG_{iso} for the remainder of this paper. We also conclude from the photometry simulations that despite the varying PSF size, MAG_{iso} is adequate because the systematics in galaxy colors averaged over all types of galaxies are less than the magnitude zero-point errors as described below.

The DLS photometric calibration was determined by observations of standard stars in Landolt's (1992) fields during photometric nights. The R and V DLS bands are very similar to Landolt's filter transmissions and yield calibration accurate to 0.02 mag based on repeated observations. The DLS B-band, however, differs significantly from Landolt's and requires a color-term correction which decreases the accuracy of calibration in this band. Also, the DLS z-band photometry derived from Sloan Digital Sky Survey standards (Smith et al. 2002) is not as good as R and V. For these reasons the B and z zero points are accurate only to 0.03 mag. BPZ folds these uncertainties into its likelihood calculations.

4.2. *Template Optimization*

We use spectroscopic redshifts and DLS photometry to empirically correct the BPZ set of templates and to test our filter + instrument response knowledge with the methodology described in Ilbert et al. (2006). We find optimized templates for the El, Sbc, Scd, Im, and SB3 SEDs. The SB2 template was left unchanged because there were not enough galaxies of this type to fit a correction. The biggest modifications were found for the El SED, which shows a less strong 4000 Å break in the optimized template, and for the Sbc SED, which has a stronger 4000 Å break than in the original BPZ template (see Fig. 19). Because most of our galaxies

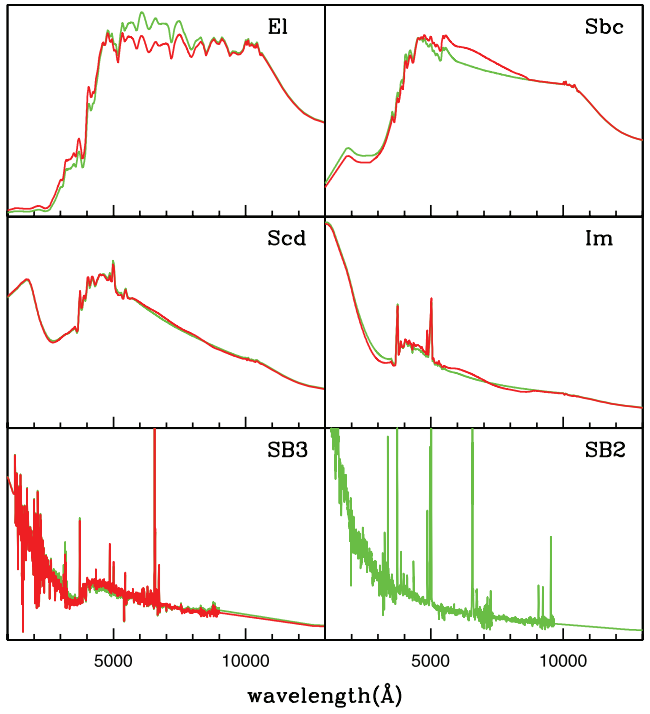

FIG. 19.—Optimized SED templates are shown in red, and the original templates are shown in green. The SB2 template was kept unchanged.

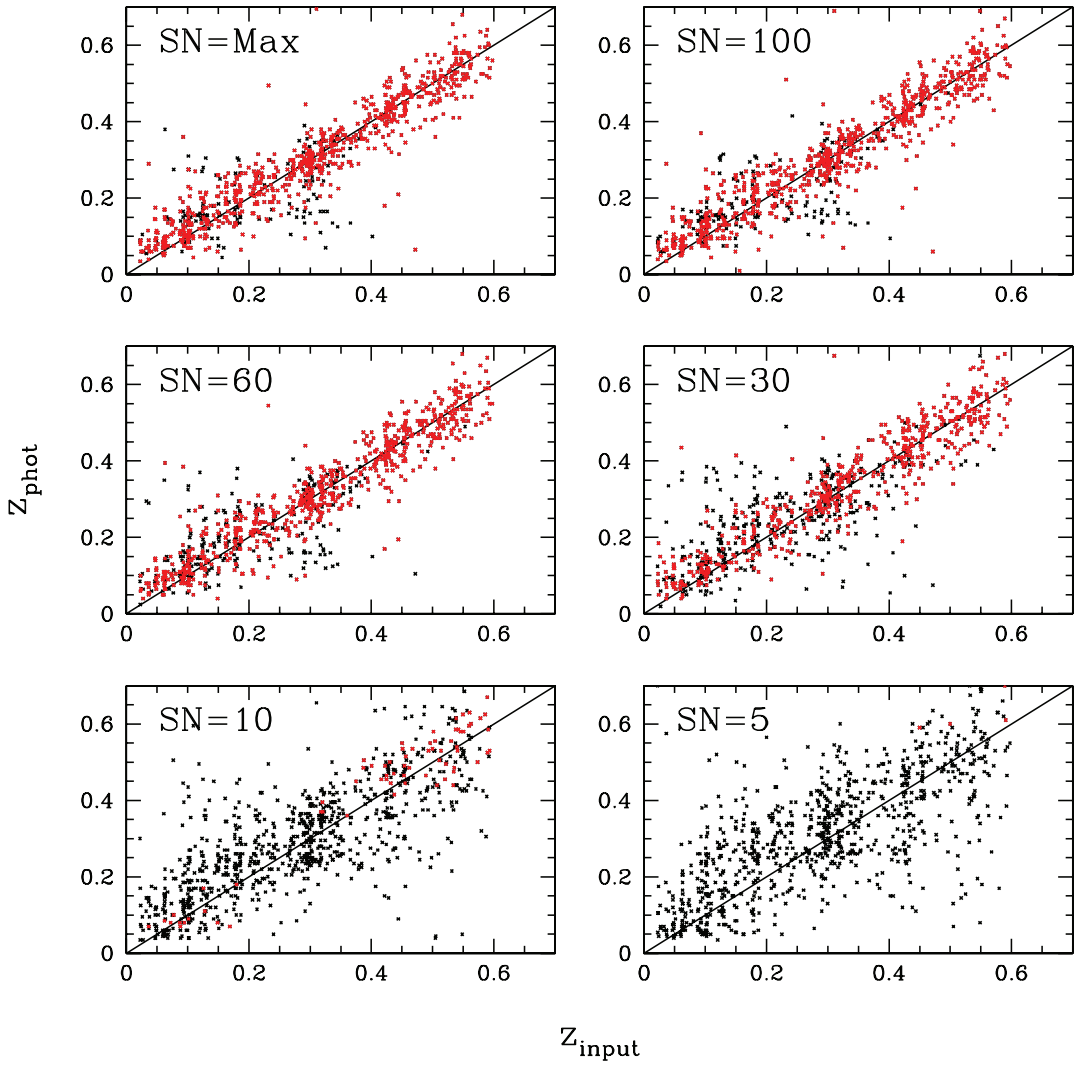

FIG. 20.— The $z_{\rm phot}$-$z_{\rm spec}$ scatter plot for 860 objects with spectroscopic redshifts from SHeLS in the DLS. The top left panel shows the results when the maximum S/N photometry is used. The other panels (*left to right, top to bottom*) show the results with progressively greater photometry noise: S/N = 100, 60, 30, 10, and 5. See Table 3 for statistics.

are at low redshift, we cannot constrain the longest and shortest SED wavelengths; therefore, we force them to agree with the initial templates.

4.3. *Comparison with Spectroscopic Data: SHeLS*

SHeLS has a limiting magnitude of $R = 20.3$, so that the DLS photometry, which is complete to about 5 mag fainter, has a very high S/N. Being a bright magnitude-limited survey, SHeLS contains overwhelmingly low-redshift ($z < 0.6$) galaxies. However, our subsample of 1000 was chosen to provide a nearly uniform redshift distribution so that characterization accuracy would be roughly redshift-independent. At a given redshift, selection was random.

We further cut the sample, requiring S/N > 100 in the R band and barring objects in exclusion zones around bright stars, with saturated pixels in any band, or with SExtractor flags ≥ 4 (compromised photometry). The final sample contains 860 galaxies. The top left panels of Figures 20 and 21 show the $z_{\rm phot}$-$z_{\rm spec}$ scatter plot and Δz distribution for the maximum S/N photometry. The distribution of galaxy types assigned by BPZ to this spectroscopic sample is in agreement with the type distribution of all galaxies at $R = 20 \pm 0.5$ mag in the entire DLS, suggesting that the spectroscopic sample is representative of galaxies at this magnitude.

The SHeLS sample is expected to show evidence of template noise and have somewhat higher $\sigma_{\Delta z}$ than the bright end of SIM3, and this is in fact observed. Objects with S/N > 100 in SIM3 have $\sigma_{\Delta z} = 0.037$, and 89.4% of the galaxies have ODDS > 0.9 with $\sigma_{\Delta z} = 0.026$. For SHeLS, $\sigma_{\Delta z} = 0.050$, and 85.6% have ODDS > 0.9 with $\sigma_{\Delta z} = 0.044$. The difference suggests a template noise of $\sigma_{\Delta z} \sim 0.035(1 + z)$, which is smaller than the $0.065(1 + z)$ estimated by Fernández-Soto et al. (1999) for galaxies in the HDF, but expected given the much lower redshift of galaxies in SHeLS.

We now degrade the photometry successively to S/N = 100, 60, 30, 10, and 5 in all bands. If a galaxy has, for example, S/N = 50 in the B band, its magnitude and magnitude error are left unchanged in this band for the simulations with S/N = 100 and 60, but noise is added to the other ones. The $z_{\rm phot}$-$z_{\rm spec}$ scatter plots are shown in Figure 20. The Δz distributions are shown in Figure 21, and the cumulative fraction as a function of Δz is shown in Figure 22. Statistics in different S/N regimes are presented in Table 3. The trends with S/N which were observed in the simulations are reproduced here.

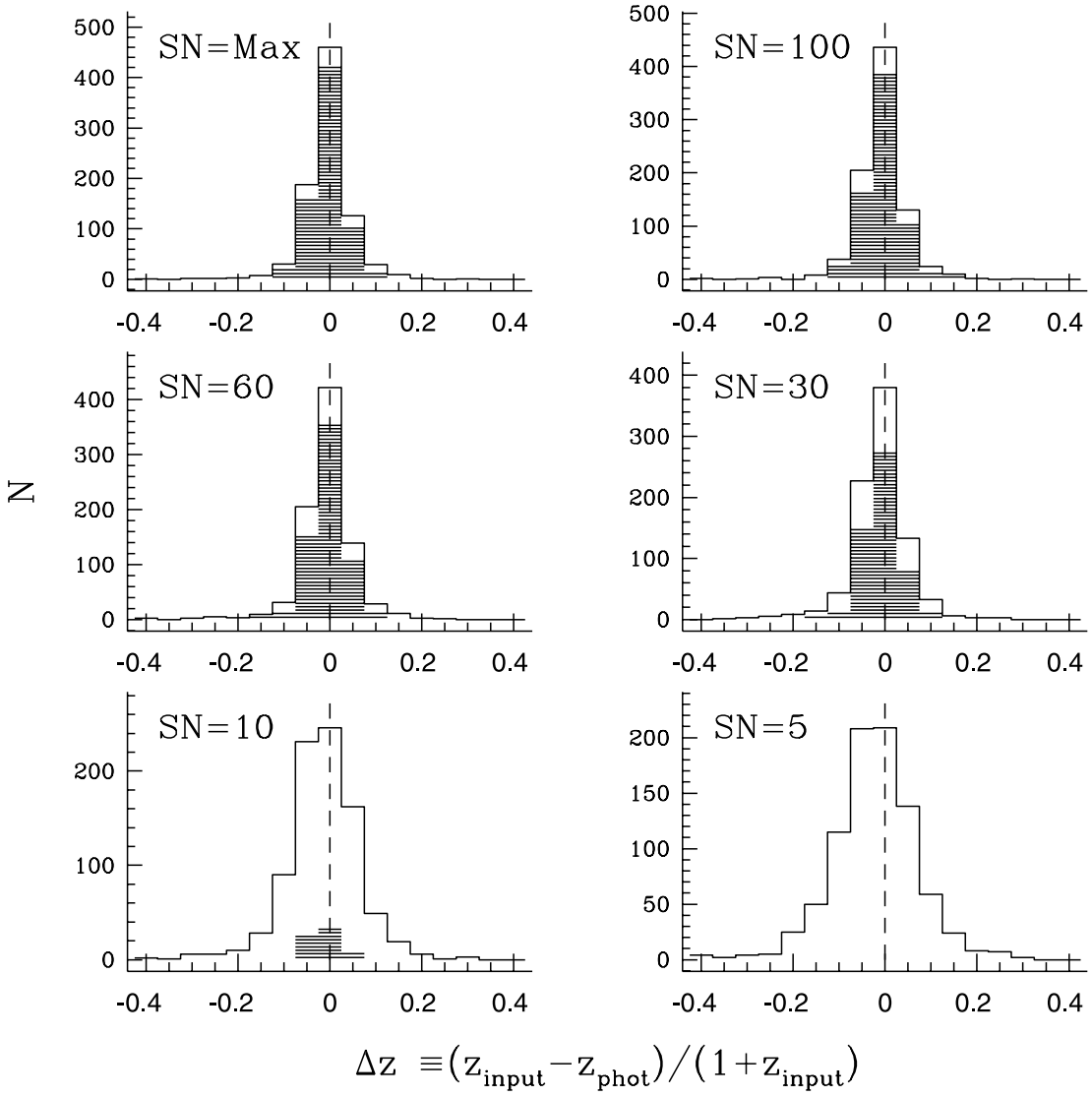

FIG. 21.—Histogram of Δz for objects shown in Fig. 20. The distribution of galaxies with ODDS > 0.9 is shaded. [*See the electronic edition of the Journal for a color version of this figure*.]

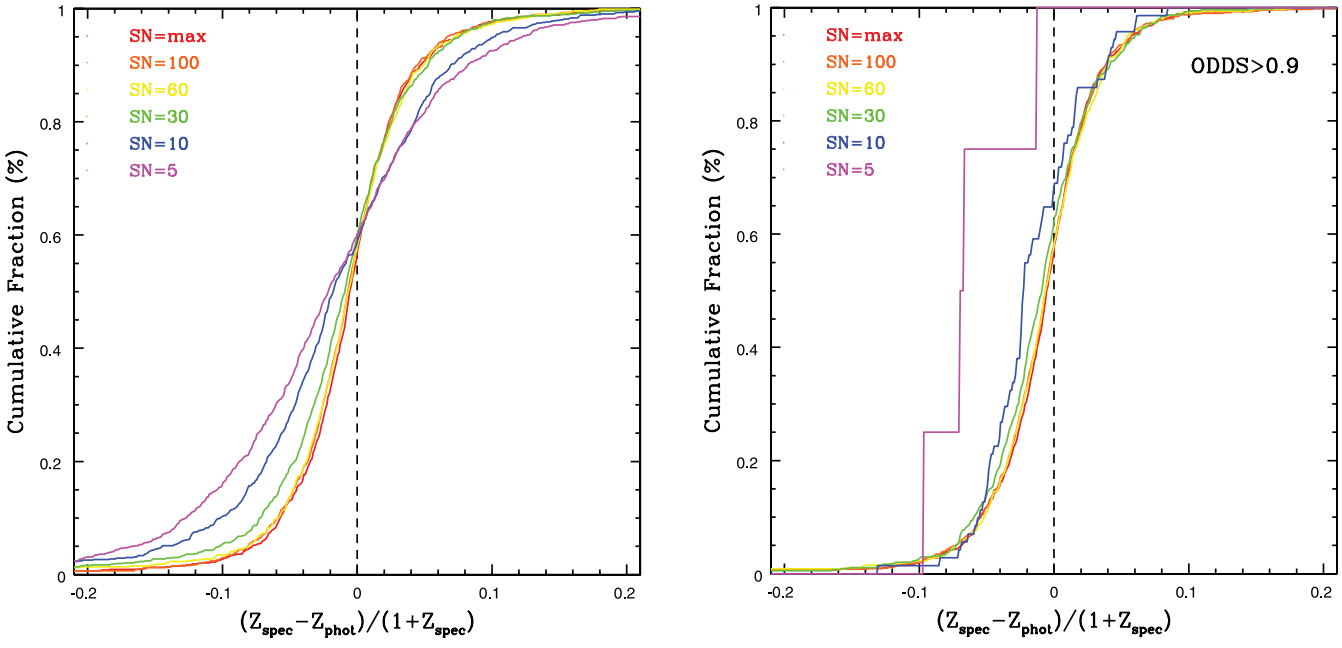

FIG. 22.—Cumulative fraction of objects as a function of Δz. The red lines show the cumulative fraction for maximum S/N photometry in SHeLS, orange shows the fraction when objects that are noised up simulate S/N = 100 in *BVRz*, and so on. The left panel shows all the galaxies, and the right panel shows galaxies with ODDS > 0.9.

TABLE 3
GALAXIES WITH SPECTROSCOPIC REDSHIFTS FROM SHeLS

	$\lvert\Delta z\rvert \leq 0.5$			ODDS > 0.9			
S/N	$\lvert\Delta z\rvert \leq 0.5$/all (%)	$\bar{\Delta} z$	$\sigma(\Delta z)$	ODDS > 0.9/all (%)	($\lvert\Delta z\rvert \leq 0.5$, ODDS > 0.9)/ODDS > 0.9 (%)	$\bar{\Delta} z$	$\sigma(\Delta z)$
Max	99.9	−0.005	0.050	85.6	99.9	−0.006	0.044
100	99.8	−0.006	0.050	83.0	99.9	−0.007	0.045
60	99.8	−0.006	0.054	76.9	99.7	−0.006	0.045
30	100.0	−0.012	0.061	62.9	100.0	−0.010	0.046
10	100.0	−0.016	0.080	8.3	100.0	−0.015	0.038
5	99.8	−0.021	0.090	0.5	100.0	−0.061	0.035

Because the magnitude prior remains tight despite the S/N degradation, we observe lower $\sigma_{\Delta z}$ at the low-S/N end of the SHeLS simulations than is observed for SIM2 at the same S/N. At S/N = 10, $\sigma_{\Delta z} = 0.080$, and 8.3% of galaxies in SHeLS have ODDS > 0.9, while $\sigma_{\Delta z} = 0.121$ and 6.4% of them have ODDS > 0.9 for the SIM2 galaxies.

The effectiveness of the ODDS cut is again evident. The fraction of galaxies passing this cut at low S/N is less than in SIM3 because the data here are uniformly at low S/N, whereas for SIM3 the given S/N is a lower limit. The fraction with ODDS > 0.9 at low S/N is more directly comparable with, and more consistent with, the fractions in SIM2, which were also at constant S/N.

The BPZ magnitude prior is specified in terms of F814W, which is close to DLS *R* but redder. In this paper we simply use *R* magnitudes in place of F814W magnitudes. To verify that this does not significantly affect the results, we computed synthetic F814W, *R*, and *z* magnitudes for all SED templates throughout the redshift range considered. We found a relatively tight (0.1 mag residuals) relation predicting F814W as a function of *R* and *z*. Using the inferred F814W magnitudes for the prior on the SHeLS data set, we found no significant improvement. This does not rule out the possibility that *R* magnitudes could be harmful in the rare special case. For example, the largest predicted discrepancy between observer frame *R* and F814W magnitudes is −1.36, for ellipticals at $z = 1.6$. These galaxies do not appear in our spectroscopic samples and therefore do not affect the results here, but they may appear in the full photometric sample. We therefore recommend using inferred F814W magnitudes for the full photometric sample.

4.4. *Comparison with Spectroscopic Data: CFGRS*

The CFGRS (Cohen et al. 1999) is about 2 mag deeper than SHeLS, and therefore the DLS photometry is not as high-S/N. We select galaxies with quality = 1 (multiple spectral features; Cohen et al. 1999) spectroscopic redshifts and divide the data into two equally sized subsamples of 111 galaxies each: one with galaxies of photometric S/N(*R*) > 106 and another with S/N(*R*) < 106. Note that the S/N in the low-S/N sample is still fairly high, with 28 being the lowest value and a median of 69, but the difference in the quality of the photometric redshifts is clear. Figure 23 shows the $z_{\rm phot}$-$z_{\rm spec}$ scatter plot for the two subsamples. For the high-S/N sample, $\Delta z = 0.027 \pm 0.084$, and $\Delta z = 0.021 \pm 0.060$ if we exclude one catastrophic outlier with $\lvert\Delta z\rvert > 0.5$. For the lower S/N sample, $\Delta z = 0.033 \pm 0.166$, and $\Delta z = 0.041 \pm 0.095$ if we exclude two objects with $\lvert\Delta z\rvert > 0.5$. However, this includes the effect of different redshift ranges. To isolate the S/N effect, we compute results using only galaxies in

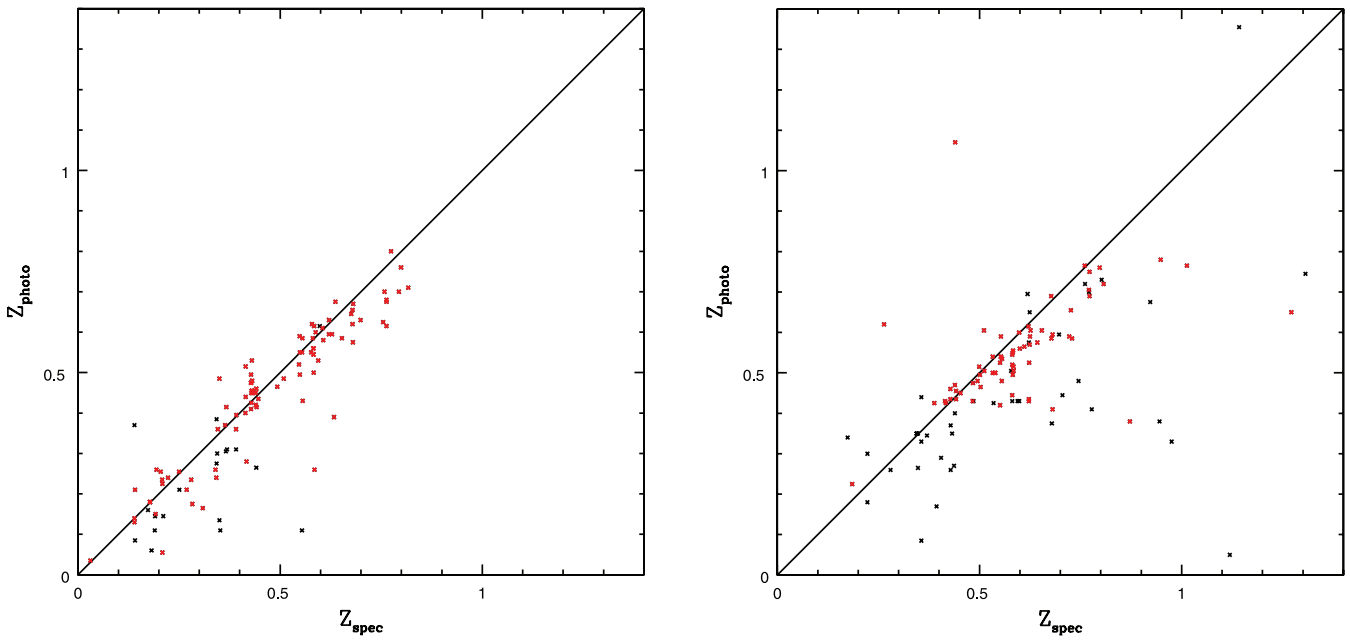

FIG. 23.— The $z_{\rm phot}$-$z_{\rm spec}$ scatter plot for 222 objects with spectroscopic redshifts from the CFGRS. The sample was subdivided: *left*, S/N > 106; *right*, S/N < 106. We show objects with ODDS > 0.9 (*red symbols*).

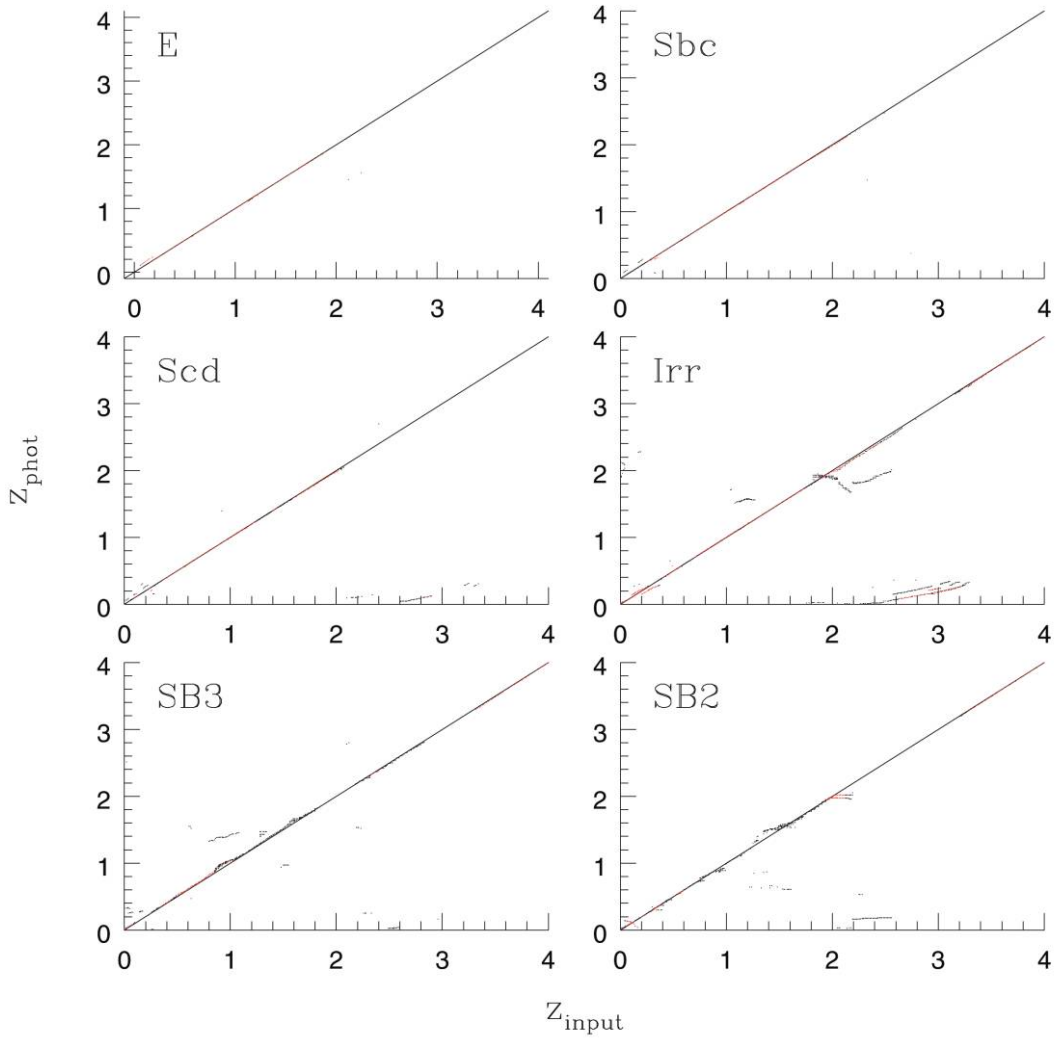

FIG. 24.—The $z_{\rm phot}$-$z_{\rm spec}$ scatter plot for SIM1 subdivided according to BPZ galaxy type (T_B). Galaxies with ODDS > 0.9 are in red.

the range $0.4 < z < 0.9$, where both samples have a significant density of sources. For the high-S/N sample we find $\Delta z = 0.020 \pm 0.056$, and for the lower S/N sample we find $\Delta z = 0.034 \pm 0.073$. No objects with $|\Delta z| > 0.5$ are found in this redshift range.

5. SELECTION IN GALAXY TYPE AND REDSHIFT RANGE

Figure 9 suggests that faint Irr/SB2/SB3 galaxies are often misclassified as Sbc/Scd. In this section we explore dependence on type in more detail. Figures 24, 25, and 26 show the $z_{\rm phot}$-$z_{\rm spec}$ scatter plot as a function of inferred BPZ galaxy type (T_B) for SIM1, SIM3, and the SHeLS galaxies, respectively. Ellipticals form the tightest relation, while the redshifts of irregular galaxies show a scatter more than twice as large. Figure 25 shows that some of the scatter in ellipticals must be due to misclassifications, because there are no E-type galaxies at $z \sim 3$–4 in the simulations.

We look at type misclassification in SIM3 directly in Figures 27 and 28. The left panels show the T_B distribution for each of the true input types, with the true type distribution overlaid like a diagonal matrix in red to guide the eye. The right panels show the true type distribution for each of the inferred types, with the inferred type distribution overlaid in red to guide the eye. The overall distribution of inferred (true) types is shown by the open histogram, which is repeated in each panel in the left (right) column. Figure 27 shows galaxies with S/N $\geq$ 30 or $R \leq 23$. For example, the fourth panel of the left column shows that galaxies classified as $T_B = 4$ (Irr) have, in fact, almost the same probability of being of types 4, 5, or 6 (irregular or starburst). Likewise, starburst galaxies tend to be misclassified as irregulars even at high S/N.

The types in decreasing order of reliability are E, Sbc, Scd, Irr, SB3, and SB2. Type reliability translates to redshift reliability, because type misclassification usually implies a large, if not catastrophic, redshift error. These figures also demonstrate that although the ODDS cut appears to lose many high-redshift galaxies and shrink the usable redshift range, in fact most of the "high-redshift" galaxies lost were type misclassifications and therefore unreliable redshifts. Although the loss of these "high-redshift" galaxies is painful if one wants as large a redshift range as possible, it is necessary if one wants the sample to be reliable.

In Figure 28 we extend the analysis to lower S/N galaxies and include all "detected" galaxies. The rate of misclassification is much higher. The insertion of these objects in the sample creates

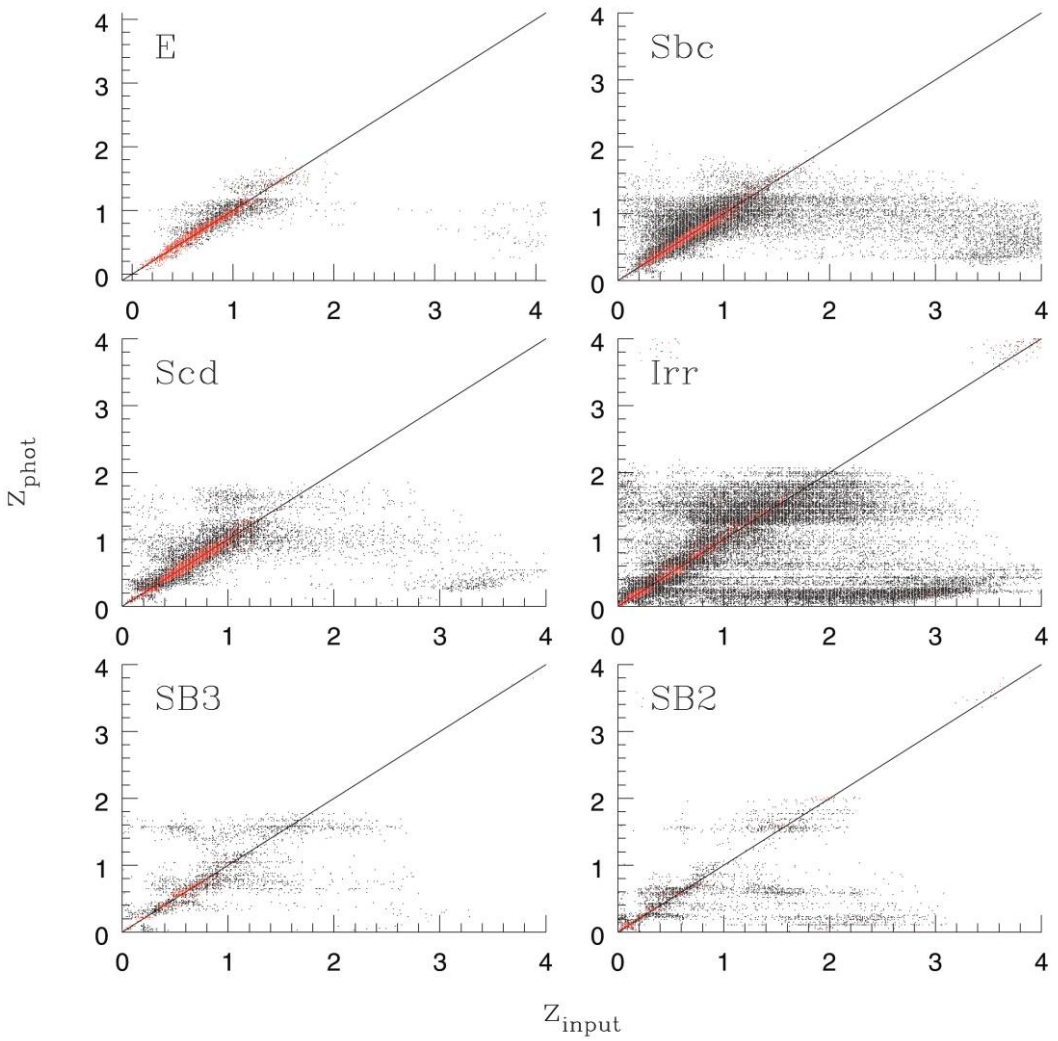

FIG. 25.—The $z_{\rm phot}$-$z_{\rm spec}$ scatter plot for SIM3 subdivided according to BPZ galaxy type (T_B). Galaxies with ODDS > 0.9 are in red.

new types of misclassification. For example, a fraction of type 1 (E) galaxies is assigned $T_B = 2$ and vice versa. Also, a significant fraction of types 4, 5, and 6 (irregular and starburst) are classified as types 2 or 3 (spirals).

6. SUMMARY AND DISCUSSION

We have examined the dependence of photometric redshift performance on photometric S/N using both simulations and data. For concreteness, we have used the DLS filter set, but the general trends should apply to any filter set. As a reminder, SIM2 simulates galaxies at a range of magnitudes drawn from the DLS photometry but at a series of constant S/N levels, while SIM3 strongly couples magnitude and S/N as they are in the DLS photometry. Thus, *bright* is distinct from *high-S/N* in SIM2 and in the noise-augmented SHeLS data because *bright* implies a more effective magnitude prior. An additional distinction between SIM3 and the other cases is that in SIM3 a given S/N cut is performed in R, and for most galaxies that implies a lower S/N in the other bands. For SIM2 and noise-augmented SHeLS data, a given S/N describes each galaxy in each filter.

We therefore expect the smallest $\sigma_{\Delta z}$ for the very high S/N in SIM3, where the high-S/N galaxies automatically have a tight magnitude prior. This is what is observed; $\sigma_z = 0.031$ (0.037) for S/N > 250 (100) in SIM3. Degeneracies in color space determine this performance limit, which is therefore highly filter set–dependent. However, it sets a baseline for what follows. At S/N $= 100$ in the SHeLS data, $\sigma_{\Delta z}$ is about 35% larger than this baseline, suggesting a cosmic variance or template noise component of $\sigma_{\Delta z} = 0.035(1 + z)$. Note that this is a shallow sample, and template noise is expected to increase for deeper samples. For SIM2, $\sigma_{\Delta z}$ is also about 32% larger than this baseline, presumably due to the looser magnitude priors on average. The deeper the survey, the less effective the magnitude prior, but performance is still quite good at this high S/N.

From this baseline, lowering the S/N smoothly increases $\sigma_{\Delta z}$ in SIM3 by 30%–50% at each S/N step in Table 2 until $\sigma_{\Delta z}$ is no longer trustworthy due to the clipping at $|\Delta z| > 0.5$. SIM2 degrades a bit more slowly due to its higher baseline $\sigma_{\Delta z}$. The noise-augmented SHeLS data degrade even more slowly because magnitude priors always remain tight. Although $\sigma_{\Delta z}$ looks reasonably good even at S/N $= 5$ for the degraded SHeLS data, we expect SIM3 to be more representative of true performance for this reason.

SIM3 indicates that without an ODDS cut, S/N $= 17$ in R is likely to be the lowest acceptable S/N for reasonable photometric

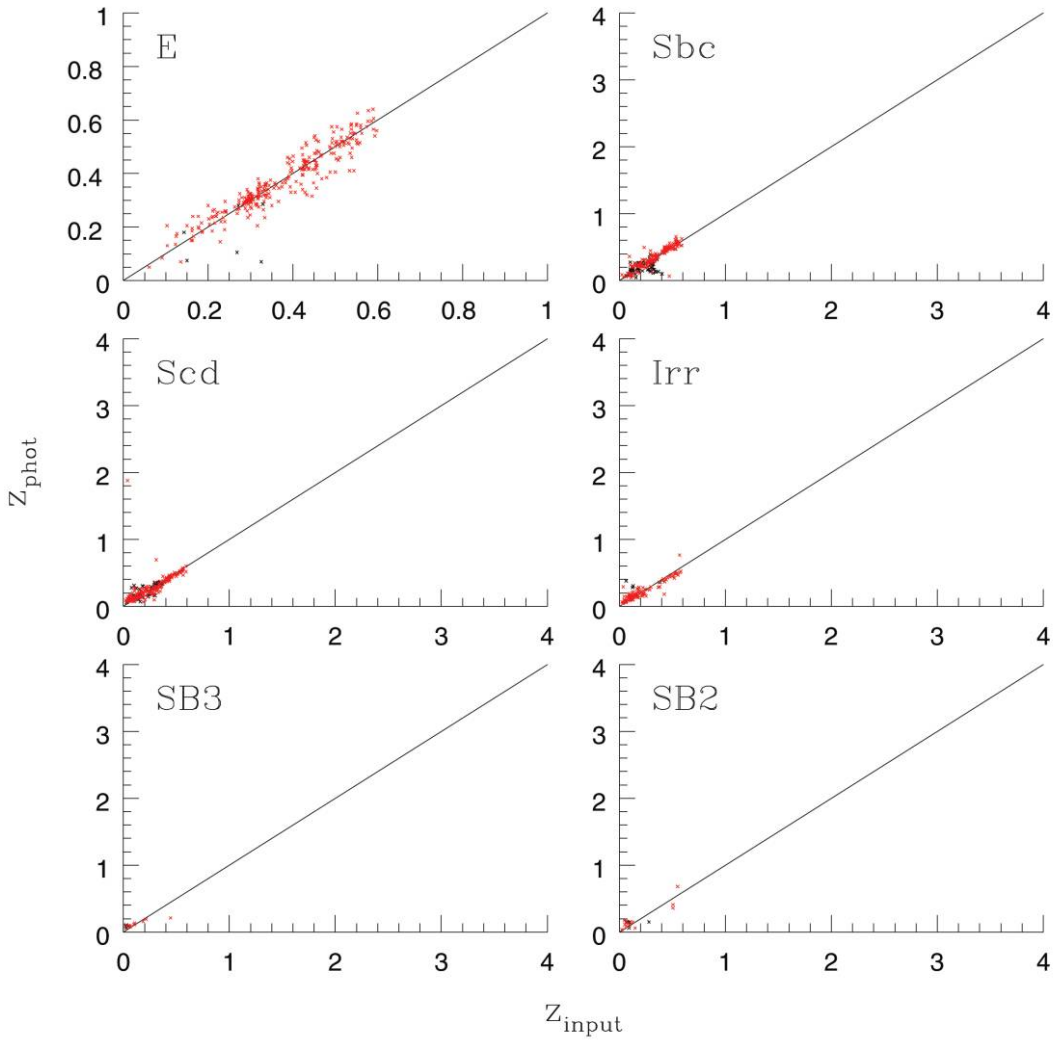

FIG. 26.—The $z_{\rm phot}$-$z_{\rm spec}$ scatter plot subdivided according to BPZ galaxy type (T_B) for 860 objects with spectroscopic redshifts from SHeLS. Galaxies with ODDS > 0.9 are in red.

redshift performance ($\sigma_{\Delta z} = 0.1$) in a survey with DLS specifications (filter set and depth). A shallower survey may be able to go to lower S/N because the magnitude prior remains helpful to lower S/N in such a survey. In fact, the bright spectroscopic sample has $\sigma_{\Delta z} < 0.1$ even at S/N $= 5$, although we caution that this means S/N $= 5$ in *each* filter. If we impose an ODDS cut rather than a S/N cut, an ODDS > 0.40 cut yields twice as many galaxies for the same $\sigma_{\Delta z}$ as the S/N > 17 cut in R. Alternatively, survey users could use ODDS to decrease $\sigma_{\Delta z}$ while sacrificing galaxy counts; an ODDS > 0.9 cut yields $\sigma_{\Delta z} = 0.04$ averaged over all S/Ns.

We caution that there are some unmodeled effects which, if included, would result in a larger $\sigma_{\Delta z}$. First, template noise is not included in the simulations. The value of $\sigma_{\Delta z}$ is larger in the SHeLS data than in SIM3 for S/N > 60, which we attribute to template noise. Template noise becomes less important at lower photometric S/N, but the template noise in the SHeLS data may be artificially low. The templates were originally derived from bright galaxies like those in SHeLS and further optimized on the SHeLS sample itself. A photometric sample which pushes to higher redshift may thus incur more template noise, and in fact Fernández-Soto et al. (1999) estimate $\sigma_{\Delta z} = 0.065(1 + z)$ for galaxies in the Hubble Deep Field.

Second, because galaxy counts are rising beyond the limiting magnitude for detection, an additional source of photometry noise must be taken into account. A source detected at a S/N of a few is much more likely to be an "upscattered" fainter galaxy than a "downscattered" brighter galaxy. As pointed out by Hogg & Turner (1998, hereafter HT98), this is distinct from Malmquist bias, which is the overrepresentation of high-luminosity galaxies in a flux-limited sample. Although the resulting bias can be computed and corrected for if the galaxy count slope is known, the additional photometric uncertainty is unavoidable. In fact, HT98 conclude that "sources identified at signal-to-noise ratios of 4 or less are practically useless." This source of noise was not reproduced in our simulations, so extrapolation to S/N < 5 would be extremely dangerous. Our results for S/N $= 5$ are still valid if 5 is interpreted as the effective S/N in the presence of this additional source of noise. For the no-evolution, Euclidean slope of $q = 1.5$, the HT98 formulae indicate that this requires a detection

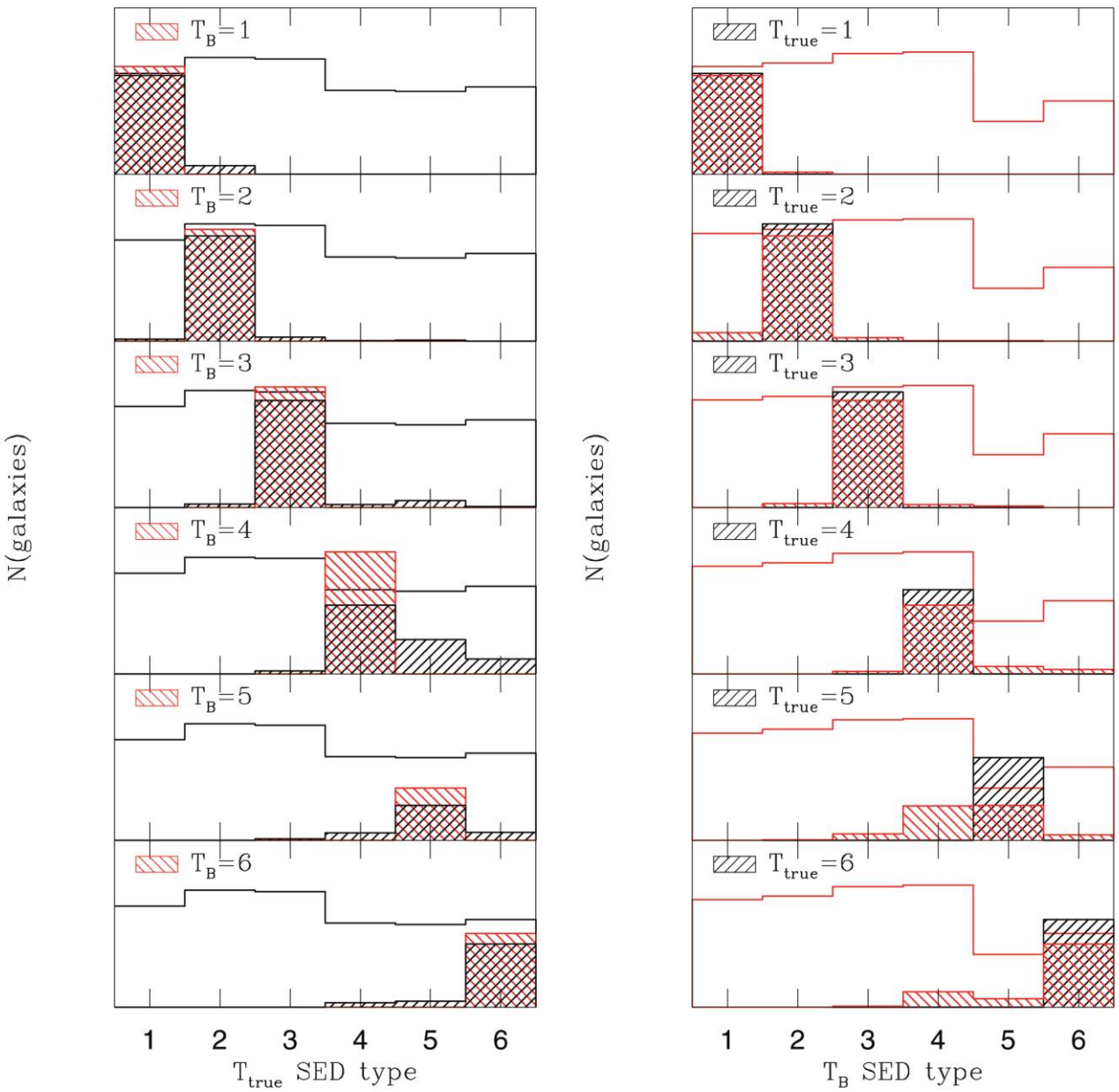

FIG. 27.—Input spectral type ($T_{\rm true}$) vs. output BPZ type (T_B) for SIM3 galaxies with S/N > 30. The types are as follows: 1, E; 2, Sbc; 3, Scd; 4, Irr; 5, SB3; and 6, SB2. In the left panels we select galaxies by T_B (*red hatched areas*) and then look at their $T_{\rm true}$ distribution (*black hatched areas*). The open black histogram is the same in all plots and indicates the $T_{\rm true}$ distribution. In the right panels, we select galaxies by their true type (*black hatched areas*) and then look at the T_B distribution (*red hatched areas*). The open red histogram indicates the T_B distribution.

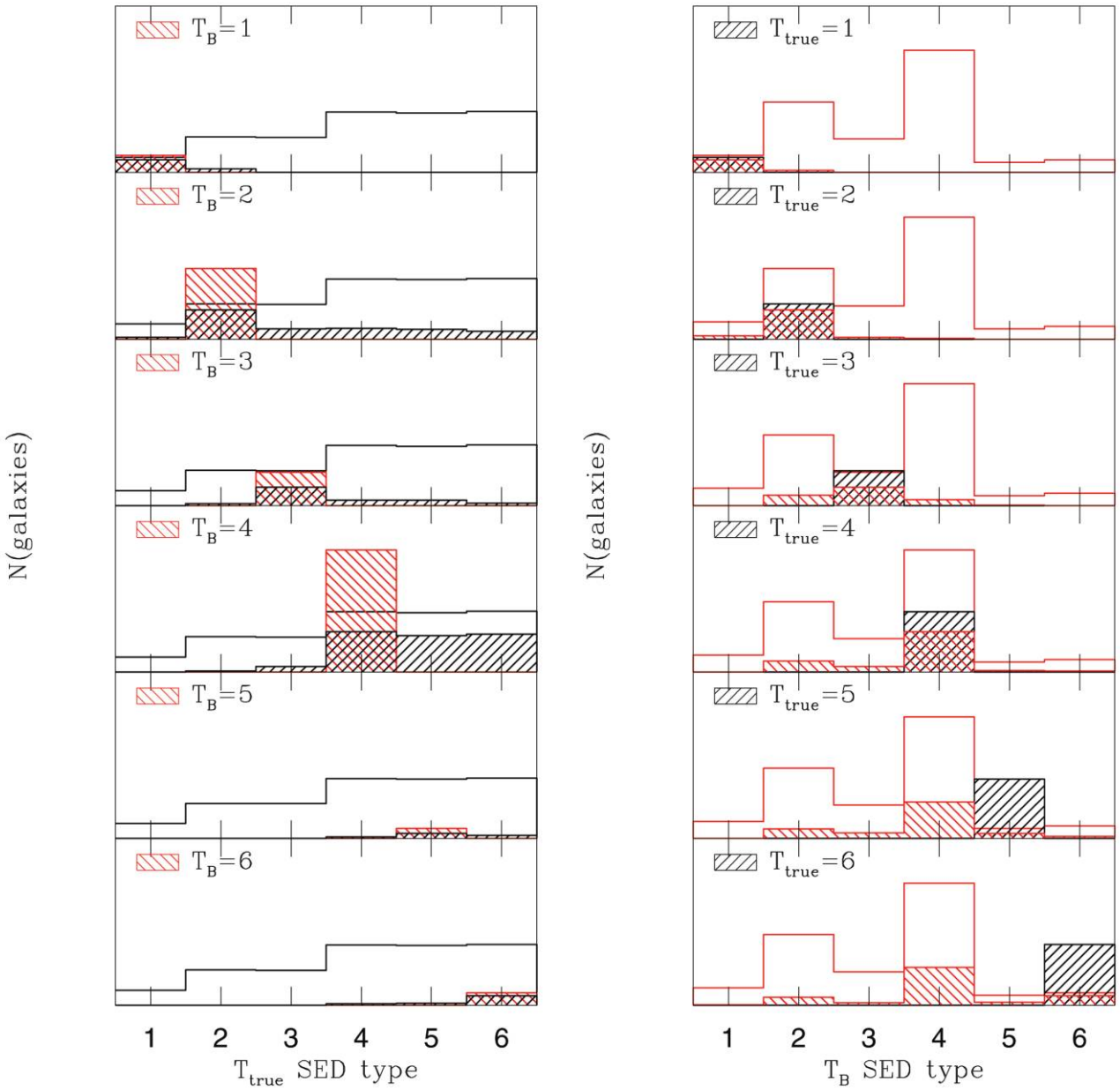

FIG. 28.—Same as Fig. 27, but including galaxies with very low S/N (all detections). Black represents true galaxy types (input), and red indicates the T_B classification.

at S/N = 5.64. For S/N = 10 and higher, the corrections are very small.

In addition to these dependences on S/N, several other lessons can be drawn:

1. When forecasting photometric redshift performance for a survey, it is important to include realistic photometry errors.

2. Estimating photometric redshift performance with spectroscopic samples can lead to optimistic results if the spectroscopic sample is not representative of the photometric sample. If the spectroscopic sample is brighter, matching the S/N is easily accomplished by adding photometry noise, but accounting for the larger redshift range of the photometric sample requires detailed modeling which must account for cosmic variance.

3. The BPZ ODDS parameter is very effective at identifying photometric redshifts which are likely to be poor. An ODDS cut is more efficient than an S/N cut, because ODDS takes account of the looser photometry requirements in distinctive regions of color space. Still, our simulations and artificially noisy data show that, of the galaxies with ODDS < 0.9, the ones with poor photometric redshifts may be in the minority. The tradeoff between the ODDS cut and the usable numbers of galaxies must be assessed in light of the specific science goal. For example, if the science analysis weights each galaxy by its photometric S/N, a strict ODDS cut may cut most of the galaxies but not most of the total weight. For weak lensing, shape noise limits the maximum weight of a galaxy, so a strict ODDS cut may cut most of the weight. Finally,

biases must be considered, as ellipticals are overrepresented in the set of galaxies with high ODDS. This may not affect weak lensing but will be important for studies of galaxy evolution and baryon acoustic oscillations.

We also explored cutting in type (as identified by BPZ) and redshift range. As expected, ellipticals do better than any other type, but we found that the ODDS cut was still useful for ellipticals. As long as the ODDS cut was being used, other types could safely be used as well. Therefore, we recommend cutting on ODDS rather than type. The ODDS cut was introduced by Benítez (2000) in the context of seven-band HDF data, including three infrared bands. It is clear from this work that it remains effective even for the much more limited *BVRz* filter set.

The DLS has received generous support from Lucent Technologies and from NSF grants AST 04-41072 and AST 01-34753. This work was also partly funded by NASA grant NNG05GD32G. DLS observations were obtained at Cerro Tololo Inter-American Observatory (CTIO) and Kitt Peak National Observatory (KPNO). CTIO and KPNO are part of the National Optical Astronomy Observatory (NOAO), which is operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation. We also would like to thank Margaret Geller and Michael Kurtz for providing us with 1000 SHeLS redshifts, which were observed with Hectospec at the MMT. We thank Ian Dell'Antonio, Tony Tyson, and the anonymous referee for comments that led to improvements to the paper.

REFERENCES

Abdalla, F. B., Amara, A., Capak, P., Cypriano, E. S., Lahav, O., & Rhodes, J. 2007, MNRAS, submitted (arXiv: 0705.1437)
Benítez, N. 2000, ApJ, 536, 571
Benítez, N., Ford, H., & Bouwens, R. 2004, ApJS, 150, 1
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476
Cohen, J. G., Hogg, D. W., Pahre, M. A., Blandford, R., Shopbell, P. L., & Richberg, K. 1999, ApJS, 120, 171
Connolly, A. J., Csabai, I., Szalay, A. S., Koo, D. C., Kron, R. G., & Munn, J. A. 1995, AJ, 110, 2655
Csabai, I., Connolly, A. J., Szalay, A. S., & Budavári, T. 2000, AJ, 119, 69
Dickinson, M. 1998, in The Hubble Deep Field, ed. M. Livio, S. M. Fall, & P. Madau (New York: Cambridge Univ. Press), 219
Fernández-Soto, A., Lanzetta, K. M., Chen, H.-W., Levine, B., & Yahata, N. 2002, MNRAS, 330, 889
Fernández-Soto, A., Lanzetta, K. M., Chen, H.-W., Pascarelle, S. M., & Yahata, N. 2001, ApJS, 135, 41
Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, ApJ, 513, 34
Geller, M. J., Dell'Antonio, I. P., Kurtz, M. J., Ramella, M., Fabricant, D. G., Caldwell, N., Tyson, J. A., & Wittman, D. 2005, ApJ, 635, L125
Hogg, D. W., & Turner, E. L. 1998, PASP, 110, 727 (HT98)
Hogg, D. W., et al. 1998, AJ, 115, 1418
Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101
Ilbert, O., et al. 2006, A&A, 457, 841
Landolt, A. U. 1992, AJ, 104, 340
Le Fèvre, O., et al. 2005, A&A, 439, 845
Ma, Z., Hu, W., & Huterer, D. 2006, ApJ, 636, 21
Smith, J. A., et al. 2002, AJ, 123, 2121
Wittman, D. M., et al. 2002, Proc. SPIE, 4836, 73
Zhan, H., & Knox, L. 2006, ApJ, 644, 663